

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 16, 2002, 14:49:43 ; Search time 23.62 Seconds
(without alignments)
482.950 Million cell updates/sec

Title: US-09-554-547-16

Perfect score: 851
Sequence: 1 MKAFVLSLTSTALTNRAR.....GSYGGSYTPNVRSAGAAA 154

Scoring table: BIOSUM62
Gapop 10.0, Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

- 1: A.GeneSeq_1101.*
- 2: /SID8/gcgdata/geneSeq/geneSeq/AA1980.DAT.*
- 3: /SID8/gcgdata/geneSeq/geneSeq/AA1981.DAT.*
- 4: /SID8/gcgdata/geneSeq/geneSeq/AA1982.DAT.*
- 5: /SID8/gcgdata/geneSeq/geneSeq/AA1983.DAT.*
- 6: /SID8/gcgdata/geneSeq/geneSeq/AA1984.DAT.*
- 7: /SID8/gcgdata/geneSeq/geneSeq/AA1985.DAT.*
- 8: /SID8/gcgdata/geneSeq/geneSeq/AA1986.DAT.*
- 9: /SID8/gcgdata/geneSeq/geneSeq/AA1987.DAT.*
- 10: /SID8/gcgdata/geneSeq/geneSeq/AA1988.DAT.*
- 11: /SID8/gcgdata/geneSeq/geneSeq/AA1989.DAT.*
- 12: /SID8/gcgdata/geneSeq/geneSeq/AA1990.DAT.*
- 13: /SID8/gcgdata/geneSeq/geneSeq/AA1991.DAT.*
- 14: /SID8/gcgdata/geneSeq/geneSeq/AA1992.DAT.*
- 15: /SID8/gcgdata/geneSeq/geneSeq/AA1993.DAT.*
- 16: /SID8/gcgdata/geneSeq/geneSeq/AA1994.DAT.*
- 17: /SID8/gcgdata/geneSeq/geneSeq/AA1995.DAT.*
- 18: /SID8/gcgdata/geneSeq/geneSeq/AA1996.DAT.*
- 19: /SID8/gcgdata/geneSeq/geneSeq/AA1997.DAT.*
- 20: /SID8/gcgdata/geneSeq/geneSeq/AA1998.DAT.*
- 21: /SID8/gcgdata/geneSeq/geneSeq/AA1999.DAT.*
- 22: /SID8/gcgdata/geneSeq/geneSeq/AA2000.DAT.*
- 23: /SID8/gcgdata/geneSeq/geneSeq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	851	100.0	154	20	AAV13500
2	180	21.2	187	21	AAAG05210
3	180	21.2	308	21	AAAG05209
4	180	21.2	358	21	AAAG05208
5	175.5	20.6	187	22	AAAB83953
6	175.5	20.6	263	21	AAAG36620
7	175.5	20.6	273	21	AAAG36619
8	175.5	20.6	309	21	AAAG36618
9	175.5	20.6	378	22	AAAG39500
10	175.5	20.6	389	22	AAAG41286
11	175	20.6	154	21	AAAG31995

12	175	20.6	256	22	AAAB30807
13	172	20.2	85	21	AAAG25162
14	172	20.2	130	21	AAAG25161
15	172	20.2	615	20	AAV40099
16	172	20.2	831	16	AAAB0168
17	171	20.1	110	21	AAAG32939
18	171	20.1	126	21	AAAB19197
19	170	20.0	136	21	AAAB19196
20	169.5	19.9	126	21	AAAG12064
21	167.5	19.7	97	21	AAAG12063
22	167.5	19.7	126	21	AAAG21911
23	166.5	19.6	147	21	AAAG25740
24	166.5	19.6	147	21	AAAG25740
25	166.5	19.6	157	21	AAAG20292
26	166	19.5	128	21	AAAB58406
27	165	19.4	175	21	AAAG24564
28	164.5	19.3	114	21	AAAG24565
29	164.5	19.3	114	21	AAAG24993
30	164	19.3	105	21	AAAG14030
31	164	19.3	147	21	AAAG14029
32	163.5	19.2	135	21	AAAG28869
33	163	19.2	132	21	AAAB19199
34	163	19.2	212	22	AAAG75147
35	163	19.2	218	21	AAAB57093
36	163	19.2	341	18	AAAB26553
37	163	19.2	341	22	AAAB81935
38	163	19.2	353	18	AAAB26552
39	163	19.2	353	19	AAAB50921
40	163	19.2	353	19	AAAB54362
41	162.5	19.1	107	21	AAAG24566
42	162.5	19.1	107	21	AAAG24994
43	161.5	19.0	133	21	AAAG42333
44	161.5	19.0	203	21	AAAG29851
45	161.5	19.0	203	22	AAAB67835

ALIGNMENTS

RESULT 1	
AAV13500	standard; Protein: 154 AA.
AAV13500;	
30-JUL-1999	(first entry)
DE	Tissue cement protein encoded by clone 64.
XX	
XX	Tissue cement protein; TCP; blood-feeding; ectoparasite; therapy;
KW	pharmaceutical; vaccine; tissue bonding; surgical wound; laceration;
KW	skin grafting; immunogen; infection; arthropod parasite.
OS	Rhipicephalus appendiculatus.
XX	
PN	W09924567-A1.
XX	
PD	20-MAY-1999.
XX	
PF	12-NOV-1998; 98WO-GB03397.
XX	
PR	12-NOV-1997; 97GB-0023945.
XX	
PA	(NAEN-) NAT ENVIRONMENTAL RES COUNCIL.
XX	
PI	Nuttall Patricia A, Paesen GC;
XX	
DR	WPI; 1999-327399/27.
DR	N-PSDB; AAX55699.
XX	
PT	Tissue cement proteins produced by blood-feeding ectoparasites and
XX	related polynucleotides

PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	21-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160980.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160989.
PR	25-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161920.
PR	28-OCT-1999;	99US-016192.
PR	28-OCT-1999;	99US-0161993.
PR	29-OCT-1999;	99US-0162142.
Query Match		
Best Local Similarity 21.2%; Score 180; DB 21; Length 187;		
Matches 56; Conservative 4; Mismatches 40; Indels 32; Gaps 10;		
Oy	33 RVHG---NLVNAIEERAGPRGYPGLTASIGCEVGARLGRAGGVSS;CYGY---PSWCY - 85	
Dd	62 fshgpaarfyag-----gygygrtgsvqpefg---ggymnyggs1g-qrnepplrys 110	
Oy	86 ----YVG-GYG----GYGAYGGYGGVDGFGSAVCGY--PGYYGYVPSPYGGGY---CG 131	
Dd	111 srfgygsgfgygegyrgygega1lyprtyggegygygygpyrgayeesgp9gsyegagg 170	
Oy	132 SYGSGSYGGSYYT 143	
Dd	171 pygrgyssrsry 182	
RESULT 3		
AAG05209	AAG05209 standard; Protein; 308 AA.	
XX	AAG05209;	
AC	AAG05209;	
XX	17-OCT-2000 (first entry)	
DT	Arabidopsis thaliana protein fragment SEQ ID NO: 1528.	
XX	Protein identification: signal transduction pathway; metabolic pathway;	
KW	hybridisation assay; genetic mapping; gene expression control; promoter;	
KW	termination sequence.	
XX	Arabidopsis thaliana.	
OS	EP1033405-A2.	
PN	06-SEP-2000.	
XX	25-FEB-2000; 2000EP-0301439.	
PF	25-FEB-1999; 99US-0121825.	
XX	05-MAR-1999; 99US-0123180.	
PR	09-MAR-1999; 99US-0123548.	
PR	23-MAR-1999; 99US-0125788.	
PR	25-MAR-1999; 99US-0126264.	
PR	29-MAR-1999; 99US-0126785.	
PR	01-APR-1999; 99US-0127462.	
PR	06-APR-1999; 99US-0128234.	
PR	08-APR-1999; 99US-0128714.	
PR	16-APR-1999; 99US-0129845.	
PR	19-APR-1999; 99US-0130077.	
PR	21-APR-1999; 99US-0130449.	
PR	23-APR-1999; 99US-0130510.	
PR	23-APR-1999; 99US-0130891.	
PR	28-APR-1999; 99US-0131449.	
PR	30-APR-1999; 99US-0132048.	

PR 30-APR-1999; 990S-0132407.
PR 04-MAY-1999; 990S-0132484.
PR 05-MAY-1999; 990S-0132485.
PR 06-MAY-1999; 990S-0132486.
PR 06-MAY-1999; 990S-0132487.
PR 07-MAY-1999; 990S-0132863.
PR 11-MAY-1999; 990S-0134256.
PR 14-MAY-1999; 990S-0134218.
PR 14-MAY-1999; 990S-0134219.
PR 14-MAY-1999; 990S-0134221.
PR 18-MAY-1999; 990S-0134370.
PR 18-MAY-1999; 990S-0134768.
PR 19-MAY-1999; 990S-0134941.
PR 20-MAY-1999; 990S-0135124.
PR 21-MAY-1999; 990S-0135353.
PR 24-MAY-1999; 990S-0135629.
PR 25-MAY-1999; 990S-0136021.
PR 27-MAY-1999; 990S-0136392.
PR 28-MAY-1999; 990S-0136782.
PR 01-JUN-1999; 990S-0137222.
PR 03-JUN-1999; 990S-0137528.
PR 04-JUN-1999; 990S-0137502.
PR 07-JUN-1999; 990S-0137724.
PR 08-JUN-1999; 990S-0138094.
PR 10-JUN-1999; 990S-0138540.
PR 14-JUN-1999; 990S-0138847.
PR 16-JUN-1999; 990S-0139119.
PR 16-JUN-1999; 990S-0139452.
PR 17-JUN-1999; 990S-0139453.
PR 17-JUN-1999; 990S-0139492.
PR 18-JUN-1999; 990S-0139454.
PR 18-JUN-1999; 990S-0139455.
PR 18-JUN-1999; 990S-0139456.
PR 18-JUN-1999; 990S-0139457.
PR 18-JUN-1999; 990S-0139458.
PR 18-JUN-1999; 990S-0139459.
PR 18-JUN-1999; 990S-0139460.
PR 18-JUN-1999; 990S-0139461.
PR 18-JUN-1999; 990S-0139462.
PR 18-JUN-1999; 990S-0139463.
PR 18-JUN-1999; 990S-0139750.
PR 18-JUN-1999; 990S-0139763.
PR 21-JUN-1999; 990S-0139817.
PR 22-JUN-1999; 990S-0139899.
PR 23-JUN-1999; 990S-0140353.
PR 23-JUN-1999; 990S-0140354.
PR 24-JUN-1999; 990S-0140695.
PR 28-JUN-1999; 990S-0140823.
PR 29-JUN-1999; 990S-0140991.
PR 30-JUN-1999; 990S-0141287.
PR 01-JUL-1999; 990S-0141842.
PR 02-JUL-1999; 990S-0142154.
PR 02-JUL-1999; 990S-0142055.
PR 06-JUL-1999; 990S-0142390.
PR 08-JUL-1999; 990S-0142803.
PR 09-JUL-1999; 990S-0142920.
PR 12-JUL-1999; 990S-0142977.
PR 13-JUL-1999; 990S-0143542.
PR 14-JUL-1999; 990S-0143624.
PR 15-JUL-1999; 990S-0144005.
PR 16-JUL-1999; 990S-0144085.
PR 16-JUL-1999; 990S-0144086.
PR 19-JUL-1999; 990S-0144325.
PR 19-JUL-1999; 990S-0144331.
PR 19-JUL-1999; 990S-0144332.
PR 19-JUL-1999; 990S-0144333.
PR 19-JUL-1999; 990S-0144334.
PR 19-JUL-1999; 990S-0144335.
PR 20-JUL-1999; 990S-0144352.
PR 20-JUL-1999; 990S-0144632.
PR 20-JUL-1999; 990S-0144884.
PR 21-JUL-1999; 990S-0144814.
PR 21-JUL-1999; 990S-0145086.

PR 21-JUL-1999; 990S-0145088.
PR 22-JUL-1999; 990S-0145085.
PR 22-JUL-1999; 990S-0145087.
PR 22-JUL-1999; 990S-0145089.
PR 22-JUL-1999; 990S-0145192.
PR 23-JUL-1999; 990S-0145145.
PR 23-JUL-1999; 990S-0145218.
PR 23-JUL-1999; 990S-0145224.
PR 26-JUL-1999; 990S-0145276.
PR 27-JUL-1999; 990S-0145913.
PR 27-JUL-1999; 990S-0145918.
PR 27-JUL-1999; 990S-0145919.
PR 28-JUL-1999; 990S-0145951.
PR 02-AUG-1999; 990S-0146386.
PR 02-AUG-1999; 990S-0146388.
PR 02-AUG-1999; 990S-0146389.
PR 03-AUG-1999; 990S-0147038.
PR 04-AUG-1999; 990S-0147204.
PR 04-AUG-1999; 990S-0147302.
PR 05-AUG-1999; 990S-0147192.
PR 05-AUG-1999; 990S-0147260.
PR 06-AUG-1999; 990S-0147303.
PR 06-AUG-1999; 990S-0147416.
PR 09-AUG-1999; 990S-0147493.
PR 09-AUG-1999; 990S-0147935.
PR 10-AUG-1999; 990S-0148171.
PR 11-AUG-1999; 990S-0148319.
PR 12-AUG-1999; 990S-0148341.
PR 13-AUG-1999; 990S-0148565.
PR 13-AUG-1999; 990S-0148684.
PR 16-AUG-1999; 990S-0149368.
PR 17-AUG-1999; 990S-0149175.
PR 18-AUG-1999; 990S-0149426.
PR 20-AUG-1999; 990S-0149722.
PR 20-AUG-1999; 990S-0149723.
PR 23-AUG-1999; 990S-0149902.
PR 23-AUG-1999; 990S-0149902.
PR 23-AUG-1999; 990S-0149930.
PR 25-AUG-1999; 990S-0150566.
PR 26-AUG-1999; 990S-0150884.
PR 27-AUG-1999; 990S-0151065.
PR 27-AUG-1999; 990S-0151066.
PR 27-AUG-1999; 990S-0151080.
PR 30-AUG-1999; 990S-0151303.
PR 31-AUG-1999; 990S-0151438.
PR 01-SEP-1999; 990S-0151930.
PR 07-SEP-1999; 990S-0152363.
PR 10-SEP-1999; 990S-0153070.
PR 13-SEP-1999; 990S-0153758.
PR 15-SEP-1999; 990S-0154018.
PR 16-SEP-1999; 990S-0154039.
PR 20-SEP-1999; 990S-0154779.
PR 22-SEP-1999; 990S-0155139.
PR 23-SEP-1999; 990S-0155486.
PR 24-SEP-1999; 990S-0155659.
PR 28-SEP-1999; 990S-0156458.
PR 29-SEP-1999; 990S-0156596.
PR 04-OCT-1999; 990S-0157117.
PR 05-OCT-1999; 990S-0157753.
PR 06-OCT-1999; 990S-0157865.
PR 07-OCT-1999; 990S-0158029.
PR 08-OCT-1999; 990S-0158232.
PR 12-OCT-1999; 990S-0158369.
PR 13-OCT-1999; 990S-0159293.
PR 13-OCT-1999; 990S-0159294.
PR 13-OCT-1999; 990S-0159295.
PR 14-OCT-1999; 990S-0159329.
PR 14-OCT-1999; 990S-0159330.
PR 14-OCT-1999; 990S-0159331.
PR 14-OCT-1999; 990S-0159637.
PR 14-OCT-1999; 990S-0159638.
PR 18-OCT-1999; 990S-0159584.
PR 21-OCT-1999; 990S-0160741.

PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 03-AUG-1999; 99US-0146389.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148341.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 11-SEP-1999; 99US-0153758.
 PR 13-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154739.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159285.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.

PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 21.2%; Score 180; DB 21; Length 358;
 Best Local Similarity 42.4%; Pred. No. 1.8e-08;

Matches 56; Conservative 4; Mismatches 40; Indels 32; Gaps 10;

QY 33 RVHG---NLVACIERAGPRGYPGLTASIGCEVGARLCGRACGVSYGYGY---PSWGY- 85
 DB 233 rshgpgasryag-----gygyrgsvvpelfg--gymnyggslyg-yrnepplygs 281
 QY 86 ----PVG-GYG----GYGAGYGAGYDGRGSAAYGY--PGYGYVYPPSGYGGY---CG 131
 DB 282 srfgpygsfgrggygrggyaylgyprggygygpyggygagsgpysyegag 341
 QY 132 SYGGSYGSGSYTY 143
 DB 342 pygrgyssary 353

RESULT 5

AA883953 standard; Protein: 187 AA.

AC AAB83953;

DT 06-AUG-2001 (first entry)

DE Amino acid sequence of the corn silk C3 gene.

XX Corn silk gene; C3 gene; C3 regulatory region; female sterile plant;

KW pathogen resistance; silk tissue; pistil tissue; stigma.

OS Zea mays.

PN CA2321269-A1.

PD 05-APR-2001.

PF 05-OCT-2000; 2000CA-2321269.

PR 05-OCT-1999; 99US-0412704.

PA (MIAC) CANADA MIN AGRIC & AGRI-FOOD CANADA.

PI Tao T, Singh J, Ouellet T;

DR WPI: 2001-329438/35.

DR N-PSDB; AAF89925.

PT New regulatory regions of corn silk for directing expression of a gene

PS of interest in stigma of both monocot and dicot plants

XX Example 3; Page 49-50; 50pp; English.
 CC The present sequence is encoded by a corn silk gene, the C3 gene. The

CC regulatory region of the C3 gene is useful for producing a female sterile
CC plant, or a plant with or improved pathogen resistance. A vector
CC comprising a chimeric gene in association with the regulatory region is
CC useful for producing a plant expressing a gene of interest within silk or
CC plastid tissue. The C3 regulatory region may be used to direct the
CC expression of a gene of interest in the stigma of both monocot and dicot
CC plants. Furthermore, the activity of the regulatory region may be
CC modified by the presence of other regulatory regions, for example,
CC enhancers, core promoter elements etc..

XX Sequence 187 AA:

Query Match 20.6% Score 175.5; DB 22; Length 187;
Best Local Similarity 38.7% Pred. No. 2.4e-08;
Matches 55; Conservative 2; Mismatches 46; Indels 39; Gaps 7;

OY 47 PRCYVCL-----TASTCGEYCARLGGAGVSSSYG-----YPSWG-----YFY 87
DB 44 papegplkdekxfgyrykhggygnmpqygggnsqpygggnsqpygy99ykrhhpg 103
OY 88 GGY-----CGYCGYGGYGY--DQFGSAVCGYPCYGYTYPSGCGCGYGSY 133
DB 104 ggygsggpggcggyggygngsgpygdnggsgtggngnag-----gy99ggyggy 158
OY 134 GSYG-GSYTYPNVRASACAA 154
DB 159 ggygsgsgtappgyghggyga 180

RESULT 6
AAC36620
ID AAC36620 standard; Protein: 263 AA.

AC AAC36620;

DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 44904.

KM Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
KM termination sequence.

XX Arabidopsis thaliana.

OS EPI033405-A2.

PN EPI033405-A2.

PD 06-SEP-2000.

XX 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-2000; 2000EP-0301439.

PR 25-FEB-1999; 9905-0121825.

PR 05-MAR-1999; 9905-0123180.

PR 09-MAR-1999; 9905-0123548.

PR 23-MAR-1999; 9905-0125788.

PR 25-MAR-1999; 9905-0126264.

PR 29-MAR-1999; 9905-0126785.

PR 01-APR-1999; 9905-0127462.

PR 08-APR-1999; 9905-0128234.

PR 16-APR-1999; 9905-0128714.

PR 19-APR-1999; 9905-0129845.

PR 21-APR-1999; 9905-0130077.

PR 23-APR-1999; 9905-0130449.

PR 28-APR-1999; 9905-0130891.

PR 30-APR-1999; 9905-0131449.

PR 04-MAY-1999; 9905-0132407.

PR 05-MAY-1999; 9905-0132484.

PR 06-MAY-1999; 9905-0132485.

PR 06-MAY-1999; 9905-0132486.

PR 06-MAY-1999; 9905-0132487.

PR 07-MAY-1999; 9905-0132863.
PR 11-MAY-1999; 9905-0134256.
PR 14-MAY-1999; 9905-0134256.
PR 14-MAY-1999; 9905-0134219.
PR 14-MAY-1999; 9905-0134221.
PR 14-MAY-1999; 9905-0134370.
PR 18-MAY-1999; 9905-0134768.
PR 19-MAY-1999; 9905-0134941.
PR 20-MAY-1999; 9905-0135124.
PR 21-MAY-1999; 9905-0135353.
PR 24-MAY-1999; 9905-0135629.
PR 25-MAY-1999; 9905-0136021.
PR 27-MAY-1999; 9905-0136392.
PR 28-MAY-1999; 9905-0136782.
PR 01-JUN-1999; 9905-0137222.
PR 03-JUN-1999; 9905-0137528.
PR 04-JUN-1999; 9905-0137502.
PR 07-JUN-1999; 9905-0137724.
PR 08-JUN-1999; 9905-0138094.
PR 10-JUN-1999; 9905-0138540.
PR 10-JUN-1999; 9905-0138847.
PR 14-JUN-1999; 9905-0139119.
PR 16-JUN-1999; 9905-0139452.
PR 16-JUN-1999; 9905-0139453.
PR 17-JUN-1999; 9905-0139452.
PR 18-JUN-1999; 9905-0139455.
PR 18-JUN-1999; 9905-0139456.
PR 18-JUN-1999; 9905-0139457.
PR 18-JUN-1999; 9905-0139458.
PR 18-JUN-1999; 9905-0139459.
PR 18-JUN-1999; 9905-0139460.
PR 18-JUN-1999; 9905-0139461.
PR 18-JUN-1999; 9905-0139462.
PR 18-JUN-1999; 9905-0139463.
PR 18-JUN-1999; 9905-0139750.
PR 18-JUN-1999; 9905-0139763.
PR 21-JUN-1999; 9905-0139817.
PR 22-JUN-1999; 9905-0139899.
PR 23-JUN-1999; 9905-0140353.
PR 23-JUN-1999; 9905-0140354.
PR 24-JUN-1999; 9905-0140695.
PR 28-JUN-1999; 9905-0140823.
PR 29-JUN-1999; 9905-0140991.
PR 30-JUN-1999; 9905-0141287.
PR 01-JUL-1999; 9905-0141842.
PR 01-JUL-1999; 9905-0142154.
PR 02-JUL-1999; 9905-0142055.
PR 06-JUL-1999; 9905-0142390.
PR 08-JUL-1999; 9905-0142803.
PR 09-JUL-1999; 9905-0142920.
PR 12-JUL-1999; 9905-0142977.
PR 13-JUL-1999; 9905-0143542.
PR 14-JUL-1999; 9905-0143624.
PR 15-JUL-1999; 9905-0144005.
PR 16-JUL-1999; 9905-0144085.
PR 16-JUL-1999; 9905-0144086.
PR 19-JUL-1999; 9905-0144325.
PR 19-JUL-1999; 9905-0144331.
PR 19-JUL-1999; 9905-0144332.
PR 19-JUL-1999; 9905-0144333.
PR 19-JUL-1999; 9905-0144334.
PR 19-JUL-1999; 9905-0144335.
PR 20-JUL-1999; 9905-0144352.
PR 20-JUL-1999; 9905-0144632.
PR 20-JUL-1999; 9905-0144634.
PR 21-JUL-1999; 9905-0144814.
PR 21-JUL-1999; 9905-0145086.
PR 21-JUL-1999; 9905-0145088.
PR 22-JUL-1999; 9905-0145085.
PR 22-JUL-1999; 9905-0145087.
PR 22-JUL-1999; 9905-0145089.
PR 22-JUL-1999; 9905-0145192.

PR	23-JUL-1999	9905-01455148
PR	23-JUL-1999	9905-01455148
PR	23-JUL-1999	9905-01455224
PR	26-JUL-1999	9905-01455276
PR	27-JUL-1999	9905-01455313
PR	27-JUL-1999	9905-01455319
PR	27-JUL-1999	9905-01455319
PR	28-JUL-1999	9905-01455519
PR	02-AUG-1999	9905-01463386
PR	02-AUG-1999	9905-01463588
PR	02-AUG-1999	9905-01465893
PR	03-AUG-1999	9905-01470308
PR	04-AUG-1999	9905-01472704
PR	04-AUG-1999	9905-01473042
PR	05-AUG-1999	9905-01471192
PR	06-AUG-1999	9905-01472603
PR	06-AUG-1999	9905-01473700
PR	06-AUG-1999	9905-01474163
PR	09-AUG-1999	9905-01478335
PR	10-AUG-1999	9905-01481711
PR	11-AUG-1999	9905-01483319
PR	12-AUG-1999	9905-01483441
PR	13-AUG-1999	9905-01485655
PR	13-AUG-1999	9905-01486884
PR	16-AUG-1999	9905-01493368
PR	17-AUG-1999	9905-01493175
PR	18-AUG-1999	9905-01494926
PR	20-AUG-1999	9905-01497222
PR	20-AUG-1999	9905-01497233
PR	20-AUG-1999	9905-01499229
PR	22-AUG-1999	9905-01499302
PR	23-AUG-1999	9905-01499330
PR	23-AUG-1999	9905-01505666
PR	26-AUG-1999	9905-01508884
PR	27-AUG-1999	9905-01510655
PR	27-AUG-1999	9905-01510866
PR	30-AUG-1999	9905-01513303
PR	31-AUG-1999	9905-01513330
PR	01-SEP-1999	9905-01513338
PR	07-SEP-1999	9905-01523263
PR	10-SEP-1999	9905-01530780
PR	13-SEP-1999	9905-01537578
PR	16-SEP-1999	9905-01540439
PR	16-SEP-1999	9905-01540339
PR	20-SEP-1999	9905-01551739
PR	22-SEP-1999	9905-01554179
PR	24-SEP-1999	9905-01554866
PR	28-SEP-1999	9905-01556539
PR	29-SEP-1999	9905-01565658
PR	04-OCT-1999	9905-01571117
PR	05-OCT-1999	9905-01575753
PR	06-OCT-1999	9905-01578655
PR	07-OCT-1999	9905-01580629
PR	08-OCT-1999	9905-01583332
PR	12-OCT-1999	9905-01585836
PR	13-OCT-1999	9905-01592933
PR	13-OCT-1999	9905-01592933
PR	14-OCT-1999	9905-01595331
PR	14-OCT-1999	9905-01596337
PR	14-OCT-1999	9905-01596338
PR	18-OCT-1999	9905-01595844
PR	21-OCT-1999	9905-01607041
PR	21-OCT-1999	9905-01607668
PR	21-OCT-1999	9905-01607770
PR	21-OCT-1999	9905-01608115

[illegible]

PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0132856.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138647.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.

PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0149932.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-015138.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.

[illegible]

DR WPI: 2001-442253/47.
DR N-PSDB: AAI58656.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX
PS Example 4: SEQ ID NO 2645; 10078bp; English.
XX
XX The invention relates to human nucleic acids (AA157798-AA161369) and
CC the encoded polypeptides (AA38642-AA42213) with neurotropic,
CC immunosuppressant and cytoskeletal activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemoclastic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
SQ Sequence 378 AA:

Query Match 20.6%; Score 175.5; DB 22; Length 378;
Best Local Similarity 37.5%; Pred. No. 4.8e-08;
Matches 57; Conservative 5; Mismatches 47; Indels 43; Gaps 10;

QY 20 RAGRLGSDLDTPGFRVHGNL-----YAGIERACGRGYPGLTASIGCEVGA 63
DB 226 rgnfifgggggnfgr-ggnfifggggggggsrsgysg-----gdggyng-----fgqd-gg 275
QY 64 RLGGAGVGVSS---YGYGPSWGCYPGYGCGYGCY-----GGYDGFSGSAYGCGP 114
DB 276 nygg--gpysrsrgygggpgpygnqgggyggyggygdynegnfgyggnynyggynyndfg 333
QY 115 GYGYVYPPSGYGGGGSYGSG-----SYGGSY 141
DB 334 nysg-qgqsnypmkgsfgrssgspgyggy 364

RESULT 10
AA41286
ID AA41286 standard; Protein: 389 AA.
XX
AC AA41286;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 6217.
XX
XX Human: neurotropic; immunosuppressant; cytoskeletal; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemoclastic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW Leukaemia.
XX
OS Homo sapiens.
XX
PN MO20015312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US34263.
XX
PR 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.

PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX
XX (HYSE-) HYSEQ INC.
XX
PI Tang Y, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
DR WPI: 2001-442253/47.
DR N-PSDB: AAI60442.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX
PS Example 2: SEQ ID NO 6217; 10078bp; English.
XX
XX The invention relates to human nucleic acids (AA157798-AA161369) and
CC the encoded polypeptides (AA38642-AA42213) with neurotropic,
CC immunosuppressant and cytoskeletal activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemoclastic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
SQ Sequence 389 AA:

Query Match 20.6%; Score 175.5; DB 22; Length 389;
Best Local Similarity 37.5%; Pred. No. 5e-08;
Matches 57; Conservative 5; Mismatches 47; Indels 43; Gaps 10;

QY 20 RAGRLGSDLDTPGFRVHGNL-----YAGIERACGRGYPGLTASIGCEVGA 63
DB 237 rgnfifgggggnfgr-ggnfifggggggggsrsgysg-----gdggyng-----fgqd-gg 286
QY 64 RLGGAGVGVSS---YGYGPSWGCYPGYGCGYGCY-----GGYDGFSGSAYGCGP 114
DB 287 nygg--gpysrsrgygggpgpygnqgggyggyggygdynegnfgyggnynyggynyndfg 344
QY 115 GYGYVYPPSGYGGGGSYGSG-----SYGGSY 141
DB 345 nysg-qgqsnypmkgsfgrssgspgyggy 375

RESULT 11
AAG31995
ID AAG31995 standard; Protein: 154 AA.
XX
AC AAG31995;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 38518.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX

PM	EP1033405-A2.
XX	06-SEP-2000.
PF	25-FEB-2000; 2000EP-0301439.
XX	
PR	25-FEB-1999; 99US-0121825.
PR	05-MAR-1999; 99US-0123180.
PR	09-MAR-1999; 99US-0123548.
PR	23-MAR-1999; 99US-0125788.
PR	25-MAR-1999; 99US-0126264.
PR	29-MAR-1999; 99US-0126785.
PR	01-APR-1999; 99US-0127462.
PR	06-APR-1999; 99US-0128234.
PR	08-APR-1999; 99US-0128714.
PR	16-APR-1999; 99US-0129845.
PR	19-APR-1999; 99US-0130077.
PR	21-APR-1999; 99US-0130449.
PR	23-APR-1999; 99US-0130510.
PR	23-APR-1999; 99US-0130891.
PR	28-APR-1999; 99US-0131449.
PR	30-APR-1999; 99US-0132048.
PR	04-MAY-1999; 99US-0132407.
PR	05-MAY-1999; 99US-0132485.
PR	06-MAY-1999; 99US-0132486.
PR	06-MAY-1999; 99US-0132487.
PR	07-MAY-1999; 99US-0132863.
PR	11-MAY-1999; 99US-0134256.
PR	14-MAY-1999; 99US-0134218.
PR	14-MAY-1999; 99US-0134219.
PR	14-MAY-1999; 99US-0134221.
PR	14-MAY-1999; 99US-0134370.
PR	18-MAY-1999; 99US-0134768.
PR	19-MAY-1999; 99US-0134941.
PR	20-MAY-1999; 99US-0135124.
PR	21-MAY-1999; 99US-0135353.
PR	24-MAY-1999; 99US-0135629.
PR	25-MAY-1999; 99US-0136021.
PR	27-MAY-1999; 99US-0136392.
PR	28-MAY-1999; 99US-0136782.
PR	01-JUN-1999; 99US-0137222.
PR	03-JUN-1999; 99US-0137528.
PR	04-JUN-1999; 99US-0137724.
PR	07-JUN-1999; 99US-0137724.
PR	08-JUN-1999; 99US-0138094.
PR	10-JUN-1999; 99US-0138540.
PR	10-JUN-1999; 99US-0138847.
PR	14-JUN-1999; 99US-0139119.
PR	16-JUN-1999; 99US-0139453.
PR	16-JUN-1999; 99US-0139453.
PR	17-JUN-1999; 99US-0139492.
PR	18-JUN-1999; 99US-0139454.
PR	18-JUN-1999; 99US-0139455.
PR	18-JUN-1999; 99US-0139456.
PR	18-JUN-1999; 99US-0139457.
PR	18-JUN-1999; 99US-0139458.
PR	18-JUN-1999; 99US-0139459.
PR	18-JUN-1999; 99US-0139460.
PR	18-JUN-1999; 99US-0139461.
PR	18-JUN-1999; 99US-0139462.
PR	18-JUN-1999; 99US-0139463.
PR	18-JUN-1999; 99US-0139464.
PR	18-JUN-1999; 99US-0139465.
PR	24-JUN-1999; 99US-0140695.
PR	28-JUN-1999; 99US-0140823.
PR	29-JUN-1999; 99US-0140991.
PR	30-JUN-1999; 99US-0141287.
PR	01-JUL-1999; 99US-0141842.
PR	01-JUL-1999; 99US-0142155.
PR	02-JUL-1999; 99US-0142055.
PR	06-JUL-1999; 99US-0142390.
PR	08-JUL-1999; 99US-0142803.
PR	09-JUL-1999; 99US-0142920.
PR	12-JUL-1999; 99US-0142977.
PR	13-JUL-1999; 99US-0143542.
PR	14-JUL-1999; 99US-0143624.
PR	15-JUL-1999; 99US-0144005.
PR	16-JUL-1999; 99US-0144085.
PR	16-JUL-1999; 99US-0144086.
PR	16-JUL-1999; 99US-0144325.
PR	20-JUL-1999; 99US-0144352.
PR	20-JUL-1999; 99US-0144632.
PR	20-JUL-1999; 99US-0144884.
PR	21-JUL-1999; 99US-0144814.
PR	21-JUL-1999; 99US-0145086.
PR	22-JUL-1999; 99US-0145087.
PR	22-JUL-1999; 99US-0145088.
PR	22-JUL-1999; 99US-0145085.
PR	22-JUL-1999; 99US-0145087.
PR	22-JUL-1999; 99US-0145089.
PR	22-JUL-1999; 99US-0145192.
PR	23-JUL-1999; 99US-0145145.
PR	23-JUL-1999; 99US-0145145.
PR	23-JUL-1999; 99US-0145148.
PR	23-JUL-1999; 99US-0145224.
PR	26-JUL-1999; 99US-0145276.
PR	27-JUL-1999; 99US-0145913.
PR	27-JUL-1999; 99US-0145918.
PR	27-JUL-1999; 99US-0145919.
PR	28-JUL-1999; 99US-0145951.
PR	02-AUG-1999; 99US-0146386.
PR	02-AUG-1999; 99US-0146388.
PR	02-AUG-1999; 99US-0146389.
PR	03-AUG-1999; 99US-0146393.
PR	03-AUG-1999; 99US-0147038.
PR	04-AUG-1999; 99US-0147204.
PR	04-AUG-1999; 99US-0147302.
PR	05-AUG-1999; 99US-0147192.
PR	05-AUG-1999; 99US-0147260.
PR	06-AUG-1999; 99US-0147303.
PR	06-AUG-1999; 99US-0147416.
PR	06-AUG-1999; 99US-0147479.
PR	09-AUG-1999; 99US-0147935.
PR	10-AUG-1999; 99

PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0128845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 23-JUN-1999; 99US-0140355.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.

PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148365.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0148684.
PR 17-AUG-1999; 99US-0149368.
PR 18-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.

PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156556.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160771.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162143.

Query Match 20.28; Score 172; DB 21; Length 85;
Best Local Similarity 45.58; Pred. No. 2.2e-08;
Matches 45; Conservative 5; Mismatches 23; Indels 26; Gaps 7;

Qy 43 ERAGRGVPGITASTIGGEVGRAGRGVSSYGYGYPMSGYRYGGYGGYGG--GYG 100
Db 8 elvpgpeyh-----ggh-----gghgagg--hyggg---ghghgghngggghldygg 50
Qy 101 GYDGFCSAYGAGPYGYGYRPSGGYGGYGGYGGYGG 139
Db 51 g---ghghghyggggghy-----ggyggggggggggggg 82

RESULT 14
AAG25161
ID AAG25161 standard; Protein; 130 AA.
XX
XX AAG25161;

XX 17-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 29112.

XX Protein identification: signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

XX Arabidopsis thaliana.

XX EPI033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128234.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 18-JUN-1999; 99US-0139817.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142820.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 16, 2002, 14:50:08 ; Search time 12.49 Seconds
(without alignments)
277.463 Million cell updates/sec

Title: US-09-554-547-16

Perfect score: 851

Sequence: 1 MKAFVLLSTALFTNAR.....GSYGSTYPNVRASGAA 154

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCTUS.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	172	20.2	832	1	US-08-209-747-2
2	172	20.2	832	1	US-08-458-298-2
3	171	20.1	126	1	US-08-963-168C-7
4	170	20.0	136	3	US-08-963-168C-6
5	163	19.2	132	3	US-08-963-168C-9
6	163	19.2	341	2	US-08-538-711A-8
7	163	19.2	341	2	US-08-725-027-8
8	163	19.2	353	4	US-08-538-711A-7
9	163	19.2	353	4	US-08-725-027-7
10	160.5	18.9	131	3	US-08-963-168C-8
11	154.5	18.2	745	2	US-09-010-928B-28
12	154.5	18.2	870	2	US-09-010-928B-2
13	153.5	18.0	116	3	US-08-963-168C-13
14	152.5	17.9	907	2	US-09-010-928B-4
15	151.5	17.8	106	3	US-08-963-168C-14
16	150.5	17.7	112	3	US-08-963-168C-16
17	146	17.2	604	4	US-08-556-978B-63
18	146	17.2	606	4	US-08-556-978B-23
19	146	17.2	606	4	US-09-247-806-6
20	146	17.2	606	4	US-09-247-806-8
21	144	16.9	111	3	US-08-963-168C-15
22	144	16.9	651	4	US-08-556-978B-19
23	144	16.9	651	4	US-09-247-806-1
24	144	16.9	718	1	US-08-425-069-2
25	144	16.9	718	2	US-08-317-844B-2
26	144	16.9	747	3	US-09-034-177-3
27	142.5	16.7	606	4	US-08-556-978B-21

28	142.5	16.7	606	4	US-09-247-806-4	Sequence 4, Appl
29	141.5	16.6	251	1	US-08-209-747-8	Sequence 8, Appl
30	141.5	16.6	251	1	US-08-458-298-8	Sequence 8, Appl
31	139.5	16.4	738	3	US-08-864-038A-3	Sequence 3, Appl
32	138	16.2	235	2	US-08-529-190B-1	Sequence 1, Appl
33	136.5	16.0	318	4	US-09-060-736-727	Sequence 727, App
34	136	16.0	334	4	US-09-060-736-728	Sequence 728, App
35	136	16.0	714	4	US-08-556-978B-61	Sequence 61, Appl
36	136	16.0	714	4	US-09-247-806-10	Sequence 10, Appl
37	134	15.7	1160	3	US-08-808-599A-24	Sequence 24, Appl
38	132	15.5	493	4	US-08-556-978B-59	Sequence 59, Appl
39	132	15.5	529	4	US-09-247-806-2	Sequence 2, Appl
40	128.5	15.1	595	1	US-08-425-069-4	Sequence 4, Appl
41	128.5	15.1	595	2	US-08-317-844B-4	Sequence 15, Appl
42	127.5	15.0	155	1	US-08-209-747-15	Sequence 15, Appl
43	127.5	15.0	155	1	US-08-458-298-15	Sequence 15, Appl
44	127	14.9	201	4	US-09-052-995-1	Sequence 1, Appl
45	127	14.9	201	4	US-09-053-003-40	Sequence 40, Appl

ALIGNMENTS

RESULT 1
US-08-209-747-2
; Sequence 2, Application US/08209747
; Patent No. 5733771
; GENERAL INFORMATION:
; APPLICANT: Lewis, Randolph V.
; APPLICANT: Colgin, Mark
; TITLE OF INVENTION: CDNAS Encoding Minor Ampullate Spider
; TITLE OF INVENTION: Silk Proteins
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch
; STREET: P.O. Box 747
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22040-3487
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/209,747
; FILING DATE: 14-MAR-1994
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Murphy Jr., Gerald M.
; REGISTRATION NUMBER: 28,977
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-205-8000
; TELEFAX: 703-205-8050
; INFORMATION FOR SEQ. ID NO. 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 832 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; FRAGMENT TYPE: Internal
; ORIGINAL SOURCE:
; ORGANISM: N. clavipes
; TISSUE TYPE: minor ampullate gland
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..309
; US-08-209-747-2

Db 1 MNAVYICLSACLIATAVSAAGYGGAGS-MGCTGCMGGMNAG-----GFGCMGGMGCG 52

QY 61 VCARLGRAGVSVSYGYGPPSMGYPGYGGYGGYGGYDGGGSAVGYPGYGY 120

Db 53 KCG-FGCGMGCGFGCMGGMG---GGP-GGFGMGCGFGCMGCGKCGCGMCGSMGCGFGM- 105

QY 121 YPSGYGGYGGYSGY-GSYCG 139

Db 106 --GGGNAGFGCMGCGCGGCG 123

RESULT 4

US-08-963-168C-6

; Sequence 6, Application US/08963168C

; Patent No. 6127166

; GENERAL INFORMATION:

; APPLICANT: Bayley, Hagan

; APPLICANT: Cao, Qunling

; APPLICANT: Wang, Yunjaun

; TITLE OF INVENTION: MOLUSCAN LIGAMENT POLYPEPTIDES

; TITLE OF INVENTION: AND GENES ENCODING THEM

; NUMBER OF SEQUENCES: 43

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish & Richardson P.C.

; STREET: 225 Franklin Street

; CITY: Boston

; STATE: MA

; COUNTRY: US

; ZIP: 02110-2804

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: Windows95

; SOFTWARE: FASTSEQ for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/963.168C

; FILING DATE: 03-NOV-1997

; ATTORNEY/AGENT INFORMATION:

; NAME: Fasse, Peter J.

; REGISTRATION NUMBER: 32,983

; REFERENCE/DOCKET NUMBER: 07917/059001

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 617/542-5070

; TELEFAX: 617/542-8906

; TELEX: 200154

; INFORMATION FOR SEQ ID NO: 6:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 136 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; FRAGMENT TYPE: internal

; FEATURE:

; OTHER INFORMATION: This translation is for SEQ ID NOS:1 & 2.

US-08-963-168C-6

Query Match 20.0%; Score 170; DB 3; Length 136;
Best Local Similarity 36.2%; Pred. No. 7.8e-09;
Matches 54; Conservative 17; Mismatches 52; Indels 26; Gaps 9;

QY 1 MKAFFVLSLSTALTNARAGRLGSDLTFCGRVHGNLYAGIERAPRGYPTLTSIGE 60

Db 1 MNAVYICLSACLIATAVSAAGYGGAGS-MGCTGCMGGMNAG-----GFGCMGGMGCG 52

QY 61 VCARLGRAGVSVSYGYGPPSMGYPGYGGYGGYGGYDGGGSAVGYPGYGY 117

Db 53 KCG-FGCGMGCGFGCMGGMG---GGP-GGFGMGCGFGCMGCGKCGCGMCGSMGCGFGM 106

QY 118 GYYPGSGYGG-----GYGGSYG-GSYCG 139

Db 107 G--GNAGFGCMGCGMAGFGCMGCGGCG 133

RESULT 5

US-08-963-168C-9

; Sequence 9, Application US/08963168C

; Patent No. 6127166

; GENERAL INFORMATION:

; APPLICANT: Bayley, Hagan

; APPLICANT: Cao, Qunling

; APPLICANT: Wang, Yunjaun

; TITLE OF INVENTION: MOLUSCAN LIGAMENT POLYPEPTIDES

; TITLE OF INVENTION: AND GENES ENCODING THEM

; NUMBER OF SEQUENCES: 43

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish & Richardson P.C.

; STREET: 225 Franklin Street

; CITY: Boston

; STATE: MA

; COUNTRY: US

; ZIP: 02110-2804

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: Windows95

; SOFTWARE: FASTSEQ for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/963.168C

; FILING DATE: 03-NOV-1997

; ATTORNEY/AGENT INFORMATION:

; NAME: Fasse, Peter J.

; REGISTRATION NUMBER: 32,983

; REFERENCE/DOCKET NUMBER: 07917/059001

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 617/542-5070

; TELEFAX: 617/542-8906

; TELEX: 200154

; INFORMATION FOR SEQ ID NO: 9:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 132 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; FRAGMENT TYPE: internal

US-08-963-168C-9

Query Match 19.2%; Score 163; DB 3; Length 132;
Best Local Similarity 35.4%; Pred. No. 3.3e-08;
Matches 52; Conservative 18; Mismatches 51; Indels 26; Gaps 8;

QY 1 MKAFFVLSLSTALTNARAGRLGSDLTFCGRVHGNLYAGIERAPRGYPTLTSIGE 60

Db 1 MNAVYICLSACLIATAVSAAGYGGAGS-MGCTGCMGGMNAG-----GFGG---TGGG 48

QY 61 VCARLGRAGVSVSYGYGPPSMGYPGYGGYGGYGGYDGGGSAVGYPGYGY 117

Db 49 MCGGKGGFGGGMGCGPGGFG--GIGGSGGCGGCGFGCMGCGKCGCGMCGSMGCGFGM 106

QY 118 GYYPGSGYGG-----GYGGSYGSGYCG 139

Db 107 G--GNAGFGCMGCGMAGFGCMGCGGCGSG--FGG 129

RESULT 6

US-08-538-711A-8

; Sequence 8, Application US/08538711A

; Patent No. 5994062

; GENERAL INFORMATION:

; APPLICANT: MUSHINE, JAMES, L.

; TITLE OF INVENTION: AN EPITHELIAL PROTEIN AND

; NUMBER OF SEQUENCES: 23

; CORRESPONDENCE ADDRESS:

TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 353
TYPE: Amino Acid
STRANDEDNESS: Unknown
TOPOLOGY: Linear
MOLECULE TYPE: Peptide
US-08-538-711A-7

Query Match 19.2%; Score 163; DB 2; Length 353;
Best Local Similarity 33.5%; Pred. No. 9.2e-08;
Matches 54; Conservative 11; Mismatches 56; Indels 40; Gaps 8;

OY 18 AARAGRLGSDLDTFGRVHGNLYAGIERAGPRGY-PGLTASI-GGEVGARLGRAGVSS 75
DB 198 SSRSGR-----GGMFGGDSRGCGGNGFGPGSGNFRGSDG--YSGRFGDGY 244
OY 76 YGYGYSMGPYGYGCGY-CGYGCGYDGFSGATGCTPGYGY- 120
DB 245 NGYGGPGGGNFGSGPGYGGGPGYGNQGGYDNYGGGNYGSGNYNDFG 304
OY 121 ----YPSGYG---GGYGS--YGSYGSYTPYNNRASAG 151
DB 305 NYNQPSNYGPKMSGNFGSRNMGGPYGGGNYGPGSGSG 345

RESULT 9

US-08-725-027-7
Sequence 7, Application US/08725027

Patent No. 6251586
GENERAL INFORMATION:
APPLICANT: MUSHINE, JAMES, L.
APPLICANT: TUCKMAN, MELVIN, S.
TITLE OF INVENTION: AN EPITHELIAL PROTEIN AND
TITLE OF INVENTION: DNA THEREOF FOR USE IN EARLY CANCER DETECTION
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/725,027
FILING DATE: 02-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US08/538,711
FILING DATE: 02-OCT-1995
ATTORNEY/AGENT INFORMATION:
NAME: KATHRYN M. BROWN
REGISTRATION NUMBER: 34,556
REFERENCE/DOCKET NUMBER: 2026-4201US1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 353
TYPE: Amino Acid
STRANDEDNESS: Unknown
TOPOLOGY: Linear
MOLECULE TYPE: Peptide
US-08-725-027-7

Query Match 19.2%; Score 163; DB 4; Length 353;
Best Local Similarity 33.5%; Pred. No. 9.2e-08;
Matches 54; Conservative 11; Mismatches 56; Indels 40; Gaps 8;

OY 18 AARAGRLGSDLDTFGRVHGNLYAGIERAGPRGY-PGLTASI-GGEVGARLGRAGVSS 75
DB 198 SSRSGR-----GGMFGGDSRGCGGNGFGPGSGNFRGSDG--YSGRFGDGY 244
OY 76 YGYGYSMGPYGYGCGY-CGYGCGYDGFSGATGCTPGYGY- 120
DB 245 NGYGGPGGGNFGSGPGYGGGPGYGNQGGYDNYGGGNYGSGNYNDFG 304
OY 121 ----YPSGYG---GGYGS--YGSYGSYTPYNNRASAG 151
DB 305 NYNQPSNYGPKMSGNFGSRNMGGPYGGGNYGPGSGSG 345

RESULT 10

US-08-963-168C-8
Sequence 8, Application US/08963168C

Patent No. 6127166
GENERAL INFORMATION:
APPLICANT: Bayley, Hagan
APPLICANT: Cao, Quidong
APPLICANT: Wang, Yunfeng
TITLE OF INVENTION: MOLLUSCAN LIGAMENT POLYPEPTIDES
TITLE OF INVENTION: AND GENES ENCODING THEM
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: US
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/963,168C
FILING DATE: 03-NOV-1997
ATTORNEY/AGENT INFORMATION:
NAME: Fasse, Peter J.
REGISTRATION NUMBER: 32,983
REFERENCE/DOCKET NUMBER: 07917/059001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 131 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: Internal
US-08-963-168C-8

Query Match 18.9%; Score 160.5; DB 3; Length 131;
Best Local Similarity 34.9%; Pred. No. 5.5e-08;
Matches 52; Conservative 20; Mismatches 46; Indels 31; Gaps 10;

OY 1 MKAFVLSLSTALTAARAGRLGSDLDTFGRVHGNLYAGIERAGPRGY-PGLTASIGE 60
DB 1 MNAVITCLAACLAIAVSAAGYGGAGS-WGCTGGGGMAG-----GG 47
OY 61 VGARLGRAGVSSYGYGPGYPSWGPYGYGCGYGYG---GGYDGFSGAVGYPGY 117

NAME: Fasse, Peter J.
REGISTRATION NUMBER: 32,983
REFERENCE/DOCKET NUMBER: 07917/059001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200134
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 116 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-963-168C-13

Query Match	18.0%	Score 153.5	DB 3	Length 116
Best Local Similarity	37.2%	Pred. No. 2.1e-07		
Matches 48	Conservative 13	Mismatches 39	Indels 29	Gaps 9

QY 21 AGRAGSLDIDFTGRVHNGNIYAGIERACGPRGYGGLTAAVSTJGEEGACILGACRACGVSSTYGY 80
 Db 4 AGSMGRT----GCMGGCGMAG-----GFGCGMGCGMGCGKG-FGGICGFGCGMGGMG- 50
 QY 81 PSMKTPPYGCGYCGYCGYCGYGGYDGGF---GSAVGYGCVYVYPSGYG-----CYGG 130
 Db 51 ---GGP--GGFGCMGCGFGCMGCGKCGFGGMSGCMGFGMGCG--GNAGFGCMGCGMAGFGG 10A
 QY 132 SYG--GYTG 139
 Db 105 MGGGCGFEG 113

RESULT 14
 US-09-010-928B-4
 : Sequence 4, Application US/09010928B
 : Patent No. 5994099
 : GENERAL INFORMATION:
 : APPLICANT: Lewis, Randolph V
 : APPLICANT: Hayashi, Cheryl Y
 : TITLE OF INVENTION: EXTREMELY ELASTIC SPIDER SILK PROTEIN AND DNA
 : TITLE OF INVENTION: CODING THEREFOR
 : NUMBER OF SEQUENCES: 29
 : CORRESPONDENCE ADDRESSES:
 : ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
 : STREET: 8110 GATEHOUSE RD. SUITE 500E
 : CITY: FALLS CHURCH
 : STATE: VIRGINIA
 : COUNTRY: UNITED STATES OF AMERICA
 : ZIP: 22042
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: PatentIn Release #1.0, Version #1.30
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/09/010.928B
 : FILING DATE: 22-JAN-1998
 : CLASSIFICATION: 435
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Murphy Jr., Gerald M
 : REGISTRATION NUMBER: 28977
 : REFERENCE/DOCKET NUMBER: 1447-109P
 : INFORMATION FOR SEQ ID NO: 4:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 907 amino acids
 : TYPE: amino acid
 : TOPOLOGY: linear
 : MOLECULE TYPE: protein
 : US-09-010-928B-4

Query match 17.98; Score 152.5; DB 2; Length 907;

Best Local Similarity 41.3%; Pred. No. 2.2e-06;
Matches 50; Conservative 11; Mismatches 39; Indels 21; Gaps 8.

[illegible]

```

RESULT 15
US-08-963-168C-14
Sequence 14, Application US/08963168C
Patent No. 6127166
GENERAL INFORMATION:
APPLICANT: Bayley, Hagan
APPLICANT: Cao, Quiding
APPLICANT: Wang, Yunlun
TITLE OF INVENTION: MOLLIUSCAN LIGAMENT POLYPEPTIDE
TITLE OF INVENTION: AND GENES ENCODING THEM
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: US
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/963,168C
FILING DATE: 03-NOV-1997
ATTORNEY/AGENT INFORMATION:
NAME: Fasse, Peter J.
REGISTRATION NUMBER: 32,983
REFERENCE/DOCKET NUMBER: 07917/059001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200134
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 106 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-963-168C-14

```

Query Match	17.8%;	Score 151.5;	DB 3;	Length 106;
Best Local Similarity	36.7%;	Pred. No. 2.9e-07;		
Matches	44;	Conservative 11;	Mismatches 44;	Indels 21; Gaps 7

OY	21	AGRAGSLDLDIFGRVHGNLYAIEERAGPRGPGYGLTASJIGEGVGAELGGRAGCGVSSYVGY	80
Dd	4	AGSMGGT-----GGMGGGMNAG-----GFGMGGMGMGGKGG-FGGMGGFSGMGGMGG-	50
OY	81	PSMGPTPGGTCGTCGTCGCTDDGPFSAISGTCFYTGTYTPSPETGGGTGGSYG-GSYGC	133
bcb	51	--GGP--GGFPGMGCPFGMGCKGKGFGGCMGSMGCFGMM---GGGNAIPFGCMGCGGGFFGG	103

Thu Jan 17 08:02:51 2002

us-09-554-547-16.ra1

Page 8

Search completed: January 16, 2002, 14:55:28
Job time: 320 sec

A:Cross-references: EMBL:U03007; NID:g927685; PIDN:AAB6485.1; PID:g927699; MIPS:YDR432w
A:Gene: SGD:NPL3; NOP3; MTS1
A:Cross-references: SGD:S0002840; MIPS:YDR432w
A:Map position: 4R
C:Superfamily: unassigned ribonucleoprotein repeat-containing proteins; ribonucleoprotein
C:Keywords: nucleolus; nucleus; RNA binding
F:126-185/Domain: ribonucleoprotein repeat homology <RRM1>
F:201-265/Domain: ribonucleoprotein repeat homology <RRM2>

Query Match 22.4% Score 190.5; DB 2; Length 414;
Best Local Similarity 45.1%; Pred. No. 1.1e-07;
Matches 55; Conservative 4; Mismatches 46; Indels 17; Gaps 7;

OY 42 IERAGPGRYGTLASTIGCEGARLG--RAGGVSSYGYPSPMGYPYGY--GGYGY- 96
 | : | : | : | : | : | : | : | : | : | : | : | : |
DB 279 IRNSNRGCFRGGRGFRGCGFGRCGFCGRCGFGCGFGCGGCGG- YGYSRGYGGGS 336
 C:Date: 02-Jul-1996 #sequence.revision 02-Jul-1996 #text.change 29-Oct-1999
 C:Accession: I37984.S40307; S77921; S41161; B35494.I37943
OY 97 -GGYGYDGGFSGAYGYPGYVPYPSGYG--GYG--GSYGGSYGGSYYTPNVRAS 149
 ||||| : | : | : | : | : | : | : | : | : | : | : | : |
DB 337 RGGYGSRRGVDSPRGGYDSPRGYSRGGYSGPRNDYPPRGSGYSGSRGC---YDPRGD 393
OY 150 AG 151
 |
DB 394 YG 395

RESULT 6
I37984
keratin 9, type I, cytoskeletal - human
N:Alternate names: cytokeratin 9; scatter protein 60K chain, placental
C:Species: Homo sapiens (man)
C:Date: 02-Jul-1996 #sequence.revision 02-Jul-1996 #text.change 29-Oct-1999
C:Accession: I37984.S40307; S77921; S41161; B35494.I37943
R:Reis, A.; Henkes, H.C.; Langbein, L.; Digeed, M.; Mitschke, D.; Drechsler, M.; Schrood
Nature Genet. 6, 174-179, 1994
A>Title: Keratin 9 gene mutations in epidermolytic palmoplantar keratoderma (EPPK).
A:Reference number: I37984; WUID:94214498
A:Accession: I37984
A>Status: preliminary; translated from GR/EMBL/DDBT
A:Molecule type: DNA
A:Residues: 1-622 <RES>
A:Cross-references: EMBL:X75015; NID:9453154; PIDN:CAAS2924.1; PID:9453155
R:Langbein, L.; Held, H.W.; Moll, I.; Franke, W.W.
Differentiation 55, 57-71, 1993
A>Title: Molecular characterization of the body site-specific human epidermal cytokerati
A:Reference number: I37943; WUID:94131202
A:Accession: S40307
A:Molecule type: mRNA
A:Residues: 1-622 <LAN>
A:Cross-references: EMBL:I229074; NID:9435475; PIDN:CAAB2315.1; PID:9435476
A:Accession: S77921
A:Molecule type: protein
A:Residues: 29-53;62-104;167-188;199-233;241-249;295-362;374-430;450-480;579-604 <LAF>
R:Langbein, L.
submitted to the EMBL Data Library, December 1993
A:Reference number: S41161
A:Accession: S41161
A:Molecule type: mRNA
A:Residues: 1-11,'SR'.13-622 <LAN>
A:Cross-references: EMBL:I229074; NID:9435475; PIDN:CAAB2315.1; PID:9435476
R:Rosen, E.M.; Merzowsky, L.; Romero, R.; Seftor, E.; Goldberg, I.
Biochem. Biophys. Res. Commun. 168, 1082-1088, 1990
A>Title: Human placenta contains an epithelial scatter protein.
A:Reference number: A35494; WUID:90267446
A:Accession: B35494
A:Molecule type: protein
A:Residues: 'X',450-465 <ROS>
C:Genetics:
A:Gene: GDB:KR19; EPPK
A:Cross-references: GDB:303970; OMIM:144200
A:Map position: 17q12-17q21

A:	Introns:	213/3:	241/2:	293/3:	347/3, 389/3: 464/2
C:	Note:	defects in this gene may cause epidermolytic palmoplantar keratoderma			
C:	Superfamily:	cytoskeletal keratin			
C:	Keywords:	coiled coil; intermediate filament			
F:	1-153/Domain:	head #status predicted <HBA>			
F:	154-459/Domain:	helical rod #status predicted <RDO>			
F:	460-622/Domain:	tail #status predicted <TAI>			
<hr/>					
Query Match	22.4%;	Score 190.5;	DB 2;	Length 622;	
Best Local Similarity	37.4%;	Pred. No. 1.6e-07;			
Matches	52;	Conservative	18;	Mismatches 42;	Indels 27; Gaps 7.
OY	11	SYPALITNARAGRLGSDLTFGRVHGNIYAGIERAPRGYPCGLTASIGSEVG--ARLGR	68		
Db	7	SSSYLTLS---GGGGGGGLSGSGSIRSS-YSRPSSSCGRGGGRFSSSSGYGGSSRVCGR	62		
OY	69	AGVGVSVSYGTGPMSKVPFYGGYGCGY-----GGYGTDDGFSAHYGCYGYCYTP	122		
Db	63	GSGG-----SFQYSYTG-G-SGGGFSASSLGGEGGGSGRCFGASGG----GYSSS	107		
OY	123	SGYGGYGGSGSYGSGSY 141			
Db	108	GGFGGGFGGGSGGGFGGCGY 126			
<hr/>					
RESULT	7				
SA0777					
heterogeneous ribonucleic particle protein A3 - African clawed frog					
N:Alternate names:	heterogeneous nuclear ribonucleoprotein A3				
C:Species:	Xenopus laevis (African clawed frog)				
C>Date:	31-Dec-1993 #sequence_revision 02-Aug-1994 #text_change 23-Jul-1999				
C:Accession:	SA0777; SA1738				
R:Good, P.J.; Rebberdt, M.L.; David, I.B.					
Nucleic Acids Res.	21, 999-1006, 1993				
A>Title:	Three new members of the RNP protein family in Xenopus.				
A:Reference number:	SA0774; MUID:93197168				
A:Accession:	SA0777				
A>Status:	nucleic acid sequence not shown				
A:Molecule type:	mRNA				
A:Residues:	1-373 <GOO1>				
A:Cross-references:	EMBL:L02956				
R:Good, P.J.; Lai, M.; Rebberdt, M.L.; David, I.B.					
Submitted to the EMBL Data Library,	October 1992				
A:Reference number:	SA1738				
A:Accession:	SA1738				
A:Molecule type:	mRNA				
A:Residues:	1-85,'S','87-346,'S','348-373 <GOO2>				
C:Cross-references:	EMBL:L02956; NID:g21474; PIDN:AAA9949.1; PID:g214745				
C:Superfamily:	helix-destabilizing protein; ribonucleoprotein repeat homology F.28-94/Domain: ribonucleoprotein repeat homology <RRM1>				
F.119-185/Domain:	ribonucleoprotein Repeat homology <RRM2>				
<hr/>					
Query Match	22.0%;	Score 187;	DB 2;	Length 373;	
Best Local Similarity	32.8%;	Pred. No. 1.9e-07;			
Matches	57;	Conservative	8;	Mismatches 55;	Indels 54; Gaps 6.
OY	16	TMAARAGRLGSLDLPFRVHGNIYAGIERAG-PRTGPGILTASIGEVARLRGARGVGS	74		
Db	201	TASAQRGGSGGSGNFMR--GGNYGGGDGNFGRGGGFGFNRRGGGCG--GRRGGGCGY	256		
OY	75	SYGYGPSMGYRYGGG-----CYGTYGTYGTYDGFSAHYGCYPRY-----	116		
Db	257	GGSDGNGFGGDCGNGYGGPGYVGRGYGSPGYGNCGGGYGGGGGYYDYDNESGNFGG	316		
OY	117	---YGY-----YPSGYGGGSGSYGSGY 139			
Db	317	YNDPFGYGGQQDSNYCPMKGSGFSGRSSSGRSGRPYGGGYSGGGGGGGSGYTG	370		
RESULT	8				
BA1732					

heterogeneous nuclear RNP protein clone pHRP40.2 - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: B41732
R:Matunis, E.L.; Matunis, M.J.; Dreyfuss, G.
J. Cell Biol. 116, 257-269, 1992
A:Title: Characterization of the major hnRNP proteins from Drosophila melanogaster.
A:Reference number: A41732; MUID:92112968
A:Accession: B41732
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-345 <MAT>
A:Cross-references: GB:X62638; NID:q11039; PIDN:CAA4504.1; PID:q11040
A>Note: Sequence extracted from NCBI Backbone (NCBIN:76632, NCBIPI:76633)
C:Genetics:
A:Gene: FlyBase:sgd
A:Cross-references: FlyBase:FBgn0003498
C:Superfamily: helix-destabilizing protein; ribonucleoprotein repeat homology
F:57-123/Domain: ribonucleoprotein repeat homology <RRM1>
F:137-203/Domain: ribonucleoprotein repeat homology <RRM2>

Query Match 21.6%; Score 183.5; DB 1; Length 345;
Best Local Similarity 45.6%; Pred. No. 3.2e-07;
Matches 52; Conservative 5; Mismatches 30; Indels 27; Gaps 9;

QY 37 NLVAGIERAGPRGYPGLTASIGEGVARGAGV----GVSSY-GYGYPSWGYPGY 90
Db 217 NQMMGMMGCGPRG--GMNGCRG-----YGRGGYNNQMDQSGVSGG-----GCG 261
QY 91 GGY--GGYGT--GGYDGFSGAYGYPGYGYPGSGY--GGYGGSYGSGSYG 139
Db 262 GGYAGGYGDIYAGGYNGYDYGDY--GYGCGFEGMNGYGGGGGGMNGCRG 314

RESULT 9
S01820
glycine-rich cell wall protein 1.8 precursor - kidney bean
C:Species: Phaseolus vulgaris (kidney bean)
C>Date: 30-Sep-1989 #sequence_revision 19-May-1994 #text_change 16-Jul-1999
C:Accession: S01820
R:Keller, B.; Sauer, N.; Lamb, C.J.
EMBO J. 7, 3625-3633, 1988
A:Title: Glycine-rich cell wall proteins in bean: gene structure and association of the
A:Reference number: S01820; MUID:89091109
A:Accession: S01820
A:Molecule type: DNA
A:Residues: 1-465 <KEL>
A:Cross-references: EMBL:X13596; NID:q21002; PIDN:CAA31932.1; PID:q21003
C:Comment: This protein is enriched in the cell wall fraction of young hypocotyls and ov
tlys.
C:Comment: Much of the sequence consists of tandemly repeated 22-residue segments with
C:Superfamily: Phaseolus glycine-rich cell wall protein 1.8
C:Keywords: cell wall; structural protein; tandem repeat
F:1-30/Domain: signal sequence #status predicted <SIG>
F:31-465/Product: glycine-rich cell wall protein 1.8 #status predicted <MAT>

Query Match 21.5%; Score 183; DB 1; Length 465;
Best Local Similarity 39.4%; Pred. No. 4.6e-07;
Matches 52; Conservative 7; Mismatches 59; Indels 14; Gaps 4;

QY 20 RAGRLSDLPFGHNLVAGIERAGPRGYPGLTASIGEGVARGAGVSSYGYG 79
Db 323 QGGGAGGGAAYVGE-HGGGYGGGGGGGGGGGGGGGGGGGGGGGAGG---YGTG 378
QY 80 YPSNGYPGYGGYGGY-----GYGYDGFSGAYGY--PGYGYPPSGYGGCY 129
Db 379 GEHGGYGG 438
QY 130 GGSYGGSYGSGY 141
Db 439 GYGAGGAHGGGY 450

RESULT 10
KNRZG1
glycine-rich cell wall structural protein 1 precursor (clone lambda-313) - rice
C:Species: Oryza sativa (rice)
C>Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 16-Jul-1999
C:Accession: S13385
R:Lei, M.; Wu, R.
Plant Mol. Biol. 16, 187-198, 1991
A:Title: A novel glycine-rich cell wall protein gene in rice.
A:Reference number: S13385; MUID:91370862
A:Accession: S13385
A:Molecule type: DNA
A:Residues: 1-165 <LEI>
A:Cross-references: EMBL:X53596; NID:q20246; PIDN:CAA37665.1; PID:q20247
C:Genetics:
A:Gene: grp-1
C:Superfamily: glycine-rich cell wall structural protein 1
C:Keywords: cell wall; duplication; structural protein
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-165/Product: glycine-rich cell wall structural protein 1 #status predicted <MAT>
F:30-55/Region: repeat R1
F:56-62/Region: repeat R2
F:62-92/Region: repeat R1
F:93-99/Region: repeat R2
F:100-131/Region: repeat R1
F:132-138/Region: repeat R2
F:139-160/Region: repeat R1

Query Match 21.2%; Score 180.5; DB 1; Length 165;
Best Local Similarity 37.5%; Pred. No. 2.9e-07;
Matches 60; Conservative 10; Mismatches 55; Indels 35; Gaps 6;

QY 5 FVLSLSTALITNAR-----ACRLGSDLPFGHNLV--AGIERA 45
Db 10 FLILLTSLSKSNARVIRKNGSGGGGGGGGGGGGGGGGGGGGGGGGGG 69
QY 46 GPRGYPGLTASIGEGVARGAGVSSYGYG-----PSWGYPGYGGYGGY 100
Db 70 GGGGSGG-----GGGGGGGGGSGSG--SGYGYGGGNGGAGCGGSGGGGGGGGG 123
QY 101 GYDGFSGAYGYPGYGYPGSGYGGGSGSYGSGS 140
Db 124 GSGGSGSGGYG-----YGTGKGGGGGGGGGGGGGGGGGSGS 158

RESULT 11
KRXL
keratin 3, type I, cytoskeletal (clone pUF1001) - African clawed frog
N:Alternate names: cytokeratin, 51K
C:Species: Xenopus laevis (African clawed frog)
C>Date: 28-Aug-1985 #sequence_revision 30-Sep-1991 #text_change 10-Dec-1999
C:Accession: S01327; A02943
R:Hoffmann, W.; Sterrer, S.; Koenigsdorfer, A.
FEBS Lett. 237, 178-182, 1988
A:Title: Amino acid sequence microheterogeneities of a type I cytokeratin of M(r) 510
A:Reference number: S01327; MUID:89005622
A:Accession: S01327
A:Molecule type: mRNA
A:Residues: 1-486 <HO>
A:Cross-references: EMBL:Y00968; NID:g64490; PIDN:CAA68783.1; PID:g64491
R:Hoffmann, W.; Praet, J.K.
EMBO J. 3, 1301-1306, 1984
A:Title: Amino acid sequence of the carboxy-terminal part of an acidic type I cytoker
A:Reference number: A02943; MUID:84261417
A:Molecule type: mRNA
A:Residues: 339-351, 'T', 353-486 <HO>
A:Cross-references: GB:X00629; NID:g64868; PIDN:CAA25263.1; PID:g833626
A>Note: this sequence has been revised in paper S01327
C:Superfamily: cytoskeletal keratin

I46032
nuclear DNA helicase II - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 02-Feb-2001
C:Accession: I46032; S49822
R:Zhang, S.; Maacke, H.; Grose, F.
J. Biol. Chem. 270, 16422-16427, 1995
A:Title: Molecular cloning of the gene encoding nuclear DNA helicase II. A bovine homolo
A:Reference number: I46032; MUID:95332357
A:Accession: I46032
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1287 <ZHA>
A:Cross-references: EMBL:X82829; NID:9577738; PIDN:CAA58036.1; PID:9577739
C:Genetics:
A:Gene: ndh2
C:Keywords: ATP; nucleotide binding; P-loop
F:408-415/Region: nucleotide-binding motif A (P-loop)
F:504-509/Region: nucleotide-binding motif B
F:508-511/Region: DEXH motif

Query Match 20.7%; Score 176.5; DB 2; Length 1287;
Best Local Similarity 37.9%; Pred. No. 3.5e-06;
Matches 53; Conservative 1; Mismatches 45; Indels 41; Gaps 5;
QY 37 NLYAGIERAGPRGYPGULTASIGFEVAGRLGGRAGVSSY-GGYPSMGYPYGGYGGY 95
Db 1144 NLMGTTRYGDPRPMPKMARYDNGSYRRGG-----SSYSGGGLGGLGTYGTGGYGGGG 1197
QY 96 YGGYGGYDOG-----FGSAY-----GGYPGYGGYYPSPGYGGG 128
Db 1198 YGGRGYSGGGYSNSFRSGYVGGGGGFGRLSRGGYRGMGSDYRGESGCG 1257
QY 129 YGGSYG-----GSYGSY 141
Db 1258 YRSGSGFQRGGRGCGYGGY 1277

Search completed: January 16, 2002, 14:55:49
Job time: 131 sec

Db 14 LIAGCGAAPVKTGRTLLGLGGFGSSVGLSAGIGVGGGLYSGGGGPGGYASGXP 72
 Qy 88 GGY-GGYGGYGGYDGFSGAY---GGY-----PGYGGYPSG-----YGGYGGS 132
 Db 73 GGYGGYSGYNGYCG---SGFGGYYPPGGGYSFGFHRPHYGHGCGYPPGGGSHNMGGSYCGH 130
 Qy 133 Y-----GGSYCG 139
 Db 131 YSOSQYNSNGYGGGCGG 148

RESULT 2
 RB56_HUMAN STANDARD; PRT; 592 AA.
 ID Q92804; Q92751;
 AC 20-AUG-2001 (Rel. 40, Created)
 DT 20-AUG-2001 (Rel. 40, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE TATA-BINDING PROTEIN ASSOCIATED FACTOR 2N (RNA-BINDING PROTEIN 56)
 DE TAFII168 (TAF(II)168).
 GN TAF2N OR RBP56.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (LONG AND SHORT ISOFORMS).
 RX MEDLINE=97124846; PubMed=8954779;
 RA Morohoshi F., Arai K., Takahashi E.-I., Tanigami A., Ohki M.,
 RT "Cloning and mapping of a human RBP56 gene encoding a putative RNA
 RT binding protein similar to FUS/TLS and EWS proteins.";
 RL Genomics 38:51-57(1996).
 RN [2]
 RP SEQUENCE FROM N.A. (SHORT ISOFORM), AND SEQUENCE OF 282-297 & 307-320.
 RX MEDLINE=97045110; PubMed=8890175;
 RA Bertolotti A., Lutz Y., Heard D.J., Chambon P., Tora L.,
 RT "TAF(II)168, a novel RNA/ssDNA-binding protein with homology to the
 RT pro-oncoproteins TLS/FUS and EWS is associated with both TFIID and RNA
 RT polymerase II.";
 RL EMBO J. 15:5022-5031(1996).
 RN [3]
 RP SEQUENCE FROM N.A. (LONG AND SHORT ISOFORMS).
 RX MEDLINE=99013873; PubMed=9795213;
 RA Morohoshi F., Ootsuka Y., Arai K., Ichikawa H., Mitani S.,
 RA Munakata M., Ohki M.;
 RT "Genomic structure of the human RBP56/TAFII68 and FUS/TLS genes.";
 RL Gene 221:191-198(1998).
 CC -|- FUNCTION: RNA AND ssDNA-BINDING PROTEIN THAT MAY PLAY SPECIFIC
 CC ROLES DURING TRANSCRIPTION INITIATION AT DISTINCT PROMOTERS. CAN
 CC ENTER THE PREINITIATION COMPLEX TOGETHER WITH THE RNA POLYMERASE
 CC II (POL. II).
 CC -|- SUBUNIT: BELONGS TO THE RNA POLYMERASE II (POL. II) TRANSCRIPTIONAL
 CC MULTIPROTEIN COMPLEX, TOGETHER WITH THE TATA-BINDING PROTEIN (TBP)
 CC AND OTHER TBP-ASSOCIATED FACTORS (TAF(II)S).
 CC -|- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
 CC -|- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG ISOFORM (SHOWN HERE)
 CC AND A SHORT ISOFORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -|- TISSUE SPECIFICITY: UBIQUITOUS. OBSERVED IN ALL FETAL AND ADULT
 CC TISSUES.
 CC -|- DISEASE: A FORM OF EXTRASKELETAL MIXOID CHONDROSARCOMAS (EMC) IS
 CC CHARACTERIZED BY A CHROMOSOMAL TRANSLOCATION T(9;17)(Q22;Q11) THAT
 CC INVOLVES TAF2N AND NR4A3.
 CC -|- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).
 CC -|- SIMILARITY: CONTAINS 1 RANBP-TYPE ZINC FINGER.
 CC -|- SIMILARITY: BELONGS TO THE TET FAMILY OF RNP PROTEINS.
 CC -----
 CC THIS SWISS-PROT ENTRY IS COPYRIGHT. IT IS PRODUCED THROUGH A COLLABORATION
 CC BETWEEN THE SWISS INSTITUTE OF BIOINFORMATICS AND THE EMBL OUTSTATION -
 CC THE EUROPEAN BIOINFORMATICS INSTITUTE. THERE ARE NO RESTRICTIONS ON ITS
 CC USE BY NON-PROFIT INSTITUTIONS AS LONG AS ITS CONTENT IS IN NO WAY
 CC MODIFIED AND THIS STATEMENT IS NOT REMOVED. USAGE BY AND FOR COMMERCIAL
 CC ENTITIES REQUIRES A LICENSE AGREEMENT (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL: U51334; AAC50932.1; -;
 DR EMBL: X98893; CA67398.1; -;
 DR EMBL: AB010067; BAA33811.1; -;
 DR EMBL: AB010057; BAA33811.1; JOINED.
 DR EMBL: AB010058; BAA33811.1; JOINED.
 DR EMBL: AB010059; BAA33811.1; JOINED.
 DR EMBL: AB010060; BAA33811.1; JOINED.
 DR EMBL: AB010061; BAA33811.1; JOINED.
 DR EMBL: AB010062; BAA33811.1; JOINED.
 DR EMBL: AB010063; BAA33811.1; JOINED.
 DR EMBL: AB010064; BAA33811.1; JOINED.
 DR EMBL: AB010065; BAA33811.1; JOINED.
 DR EMBL: AB010066; BAA33811.1; JOINED.
 DR EMBL: AB010067; BAA33812.1; -;
 DR EMBL: AB010057; BAA33812.1; JOINED.
 DR EMBL: AB010058; BAA33812.1; JOINED.
 DR EMBL: AB010059; BAA33812.1; JOINED.
 DR EMBL: AB010060; BAA33812.1; JOINED.
 DR EMBL: AB010061; BAA33812.1; JOINED.
 DR EMBL: AB010062; BAA33812.1; JOINED.
 DR EMBL: AB010063; BAA33812.1; JOINED.
 DR EMBL: AB010064; BAA33812.1; JOINED.
 DR EMBL: AB010065; BAA33812.1; JOINED.
 DR EMBL: AB010066; BAA33812.1; JOINED.
 DR MIM: 601574; -;
 DR InterPro: IPR000504; RRM.
 DR InterPro: IPR001876; Znf-RanBP.
 DR Pfam: PRF0076; rrm; 1.
 DR Pfam: PRF0641; znf-RanBP; 1.
 DR SMART: SM00547; ZNF_RB2; 1.
 DR SMART: SM00547; ZNF_RB2; 1.
 DR PROSITE: PS50102; RRM; 1.
 DR PROSITE: PS00030; RRM_RNP.1; FALSE NEG.
 KW Proto-oncogene: RNA-binding; DNA-binding; Nuclear protein; Repeat;
 KW Alternative splicing; Chromosomal translocation; Zinc finger; Zinc;
 KW Metal-binding.
 FT DOMAIN 1 208 GLN/GLY/SER/TYR-RICH.
 FT 234 320 RNA-BINDING (RRM).
 FT 320 590 ARG/GLY-RICH.
 FT ZN_FING 360 379 CA-TYPE (POTENTIAL).
 FT DOMAIN 407 575 21 X APPROXIMATE TANDEM REPEATS OF D-R-
 FT {S,G}(0,3)-G-G-Y-G-G.
 FT REPEAT 407 413 1.
 FT REPEAT 414 420 2.
 FT REPEAT 421 429 3.
 FT REPEAT 430 439 4.
 FT REPEAT 440 448 5.
 FT REPEAT 449 457 6.
 FT REPEAT 458 465 7.
 FT REPEAT 466 473 8.
 FT REPEAT 474 481 9.
 FT REPEAT 482 488 10.
 FT REPEAT 489 496 11.
 FT REPEAT 497 503 12.
 FT REPEAT 504 510 13.
 FT REPEAT 511 517 14.
 FT REPEAT 518 524 15.
 FT REPEAT 525 533 16.
 FT REPEAT 534 543 17.
 FT REPEAT 544 551 18.
 FT REPEAT 552 560 19.
 FT REPEAT 561 568 20.
 FT REPEAT 569 575 21.
 FT VARSPLIC 60 62 MISSING (IN SHORT ISOFORM).
 SQ SEQUENCE 592 AA; 61830 MW; 73D37C171E1E2BCA CRC64;

Qy 20 RAGRLGSDLDTPGFRVHGNLYAGIERAGPRGYPGLTASIGGEVVGARLGG-RAGVGVSSYGY 78
 Query Match 23.3%; Score 198.5; DB 1; Length 592;
 Best Local Similarity 43.5%; Pred. No. 2.6e-07;
 Matches 60; Conservative 3; Mismatches 52; Indels 23; Gaps 9;

Db 459 RGGYGGD-----RGGYGG-DGGG--GYGDDRGCGDGGGCGGCGDGGCGY 508
 QY 79 GYPSMGY--PYGGYGGGCGYGGYGGFSGAYGC--YGGYGYPPSGY-----GGGCG 130
 Db 509 GCRGCGGCGRGCGGCGDRSGRGGTGG-DGGGSGYGGDRSGGCGDRSGGCGGCGGCG 567
 QY 131 ---GSGYSGYSGSYTYPPN 145
 Db 568 GDRGCGGCKMGKGRNDYRN 585

RESULT 3

ID GRP_HORVU STANDARD: PRT: 200 AA.

AC P17816;
 DT 01-AUG-1990 (Rel. 15, Last Created)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE GLYCINE-RICH CELL WALL STRUCTURAL PROTEIN PRECURSOR.
 GN GRP.

OS Hordeum vulgare (Barley).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
 OC Triliceae; Hordeum.
 OX NCBI_TaxID=4513;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN-CV. ABRYSINIAN 2231; TISSUE-Leaf;
 RX MEDLINE=91346692; PubMed=1715208;

RA Rohde W., Rosch K., Kroegeer K., Salami F.;
 RT "Nucleotide sequence of a Hordeum vulgare gene encoding a glycine-rich
 protein with homology to vertebrate cytochromes.";
 RL Plant Mol. Biol. 14:1057-1059(1990).

CC -1- FUNCTION: RESPONSIBLE FOR PLASTICITY OF THE CELL WALL (POTENTIAL).
 CC -1- INDUCTION: UPON WOUNDING OF PLANT TISSUE.

CC -1- SIMILARITY: TO VERTEBRATE CYTOCHROMES.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: X52580; CAA36811.1; -;
 DR PIR: S10334; S10334.
 DR HISSP: P02876; 2MGC.

KW Cell wall; Structural protein; Repeat; Signal.

FT STGNL 1 24
 FT CHAIN 25 200 GLYCINE-RICH CELL WALL STRUCTURAL
 FT PROTEIN.

SO SEQUENCE 200 AA; 18108 MW; E8416CDEA6D49E13 CRC64;

Query Match 22.5%; Score 191.5; DB 1; Length 200;
 Best Local Similarity 36.5%; Pred. No. 3.2e-07;
 Matches 65; Conservative 9; Mismatches 59; Indels 45; Gaps 9;

QY 2 KAEVVLSTLSTAALTN-----ARAGRLGSDLDTF-----GRVHGLVAGIERAGRGY 50
 Db 6 KGLVLVALLLAALILVASADENHRAKKENBAGYENFHHGCGCHNGRRG---HGCGGY 62
 QY 51 PGLTASTGGEVAGRLGAGYVSSYGYPSMGP-----YGYGYG-----93
 Db 63 GG-----GGYGGGGGGYGGG---GGYGGGGGGYRGHGBEGGGYGGGGYRGHGBEGG 114
 QY 94 GGYGCGYGGYGGFSGAYGYPGYGYPYGGYGGGSGYSGSYGTYTPNVRASAG 151
 Db 115 GGTGGGGGTYGHGEGGGGGYGGGGYH---GHGEGGGGGYGGGGG---YFGHGGGG 166

RESULT 4

ID N0P3_YEAST STANDARD: PRT: 414 AA.

AC 001560;
 DT 01-JUL-1993 (Rel. 26, Last Created)

DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE NUCLEOLAR PROTEIN 3 (MITOCHONDRIAL TARGETING SUPPRESSOR 1 PROTEIN).
 GN N0P3 OR NPL3 OR YDR432W OR D9461.19.
 OS Saccharomyces cerevisiae (Baker's yeast).

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxID=4932;
 RN [1]

RP SEQUENCE FROM N.A.
 RX MEDLINE=93054908; PubMed=1429834;

RA Russell I.D., Tollervey D.;
 RT "N0P3 is an essential yeast protein which is required for pre-rRNA
 RT processing.";
 RL J. Cell Biol. 119:737-747(1992).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=93004942; PubMed=1392078;
 RA Bossie M.A., Denotatious C., Barcejo G., Silver P.;
 RT "A mutant nuclear protein with similarity to RNA binding proteins
 RT interferes with nuclear import in yeast.";
 RL Mol. Biol. Cell 3:875-893(1992).

RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94040759; PubMed=8224861;

RA Ellis E.M., Reid G.A.;
 RT "The Saccharomyces cerevisiae MTS1 gene encodes a putative
 RT RNA-binding protein involved in mitochondrial protein targeting.";
 RL Gene 132:175-183(1993).

RN [4]

RP SEQUENCE FROM N.A.

RA Dietrich F.S., Mulligan J., Allen E., Araujo R., Aviles E.,
 RA Barro A., Carpenter J., Chen E., Cherry J.M., Chung E., Duncan M.,
 RA Hunnicke-Smith S., Hyman R., Komp C., Lashkari D., Lew H., Lin D.,
 RA Moseedale D., Nakahara K., Namath A., Oefner P., Oh C., Petel F.X.,
 RA Roberts D., Schramm S., Schroeder M., Shogren T., Shroff N.,
 RA Whitall A., Yelton M., Botstein D., Davis R.W.;

CC Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: REQUIRED FOR PRE-RRNA PROCESSING AND N-LEAD IMPORT
 CC AS WELL AS MITOCHONDRIAL PROTEIN TARGETING.

CC -1- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLUS.

CC -1- SIMILARITY: BELONGS TO THE GAR FAMILY.

CC -1- SIMILARITY: CONTAINS 2 RNA RECOGNITION MOTIFS (RRM).

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: X66019; CAA46817.1; -;
 DR EMBL: M86731; AAA34818.1; -;
 DR EMBL: X70951; CAA50291.1; -;
 DR EMBL: U33007; AAB64865.1; -;
 DR PIR: A44401; A44401.
 DR PIR: S27435; S27435.
 DR PIR: S31322; S31322.
 DR PIR: JN0866; JN0866.
 DR PIR: S42161; S42161.
 DR HSSP: P19339; 2SXL.
 DR SGD: S0002840; NPL3.
 DR InterPro: IPR000504; RRM.
 DR Pfam: PF00076; rrm; 2.
 DR SMART: SM00360; RRM; 2.
 DR PROSITE: PS50102; RRM; 2.
 DR PROSITE: PS00030; RRM_NRP_1; FALSE_NEG.


```

RT      "A novel glycine-rich cell wall protein gene in rice." ;
RL      Plant Mol. Biol. 16:187-196(1991).
CC      -1- FUNCTION: RESPONSIBLE FOR PLASTICITY OF THE CELL WALL (POTENTIAL).
CC      -1- SUBCELLULAR LOCATION: CELL WALL (POTENTIAL).
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL: X53596; CAA37665.1; -.
DR      PIR: S13385; KNR2G1.
DR      HSSP: P30129; 4DPV.
KW      Cell wall; Structural protein; Repeat; Signal.
FT      SIGNAL                     1      23      POTENTIAL.
FT      CHAIN                     24      165      GLYCINE-RICH CELL WALL STRUCTURAL
FT                                     PROTEIN 1.
FT      DOMAIN                   31      159      GLY-RICH.
FT      REPEAT                   56      62      R2 (TYR-RICH) .
FT      REPEAT                   93      99      R2 (TYR-RICH) .
FT      REPEAT                   132     138      R2 (TYR-RICH) .
SQ      SEQUENCE 165 AA: 13536 MW: E36CE31C3650AC9A CXC64:
      Query Match      21.2%; Score 180.5; DB 1; Length 165;
      Best Local Similarity 37.5%; Pred. NO. 1.6e-06;
      Matches 60; Conservative 10; Mismatches 55; Indels 35; Gaps 6;
OY      5 FVLSLSTLAALTNAR-----AGRLSDLDTFGRVHGNLY--AGTERA 45
DB      10 FLLLTITSLSKNAGARIVKYNCGSGCGGCGGCGGCGGSGSGYGYGKAGQSG 69
OY      46 GPRGVPGLASTIGGEVARGLRAGAGVSSVSYGYT----PSMGVPGYGGYGGYGGY 100
DB      70 GGGGGSGG-----GGGGGGGGGNGSGSG--SGYGYGAGGNGCAQGGGSGGGGGGGGGGG 123
OY      101 GYDQGFSGSNVGGYPCGYGYYPGYPGGYGGYGGSGSYGGS 140
DB      124 GSGGGSGSGYG-----YGYKGGGGGGGGGGGGGGGGGGGS 158
RESULT 10
DDX9_MOUSE STANDARD: PRT: 1380 AA.
ID      DDX9_MOUSE
AC      070133; O35931; O54703;
DT      15-DEC-1998 (Rel. 37, Created)
DT      15-DEC-1998 (Rel. 37, Last sequence update)
DT      20-AUG-2001 (Rel. 40, Last annotation update)
DE      ATP-DEPENDENT RNA HELICASE A (NUCLEAR DNA HELICASE II) (NDH II)
DE      (DEAD-BOX PROTEIN 9) (MHE1-5).
GN      DDX9.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
OX      NCBI_Taxid=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=129/SV;
RX      MEDLINE=98149984; PubMed=9480750;
RX      Lee C.-G., Eki T., Okumura K., da Costa Soares V., Hurwitz J.;
RT      "Molecular analysis of the cdna and genomic DNA encoding mouse RNA
RT      helicase A." ;
RL      Genomics 47:365-371(1998).
RN      [2]
RN      SEQUENCE OF 386-919 FROM N.A.
RC      STRAIN=C57BL/6;
RA      Kistelow P., Miazek A.;
RL      Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
CC      -1- FUNCTION: UNKNINS DOUBLE-STRANDED DNA AND RNA IN A 3' TO 5'
CC      DIRECTION. GENERATES MULTIPLE SECONDARY STRUCTURES THAT INFLUENCE

```

CC	RNABINDING PROTEINS. MAY PLAY A ROLE IN X-LINKED GENE EXPRESSION.
CC	-I- SUBCELLULAR LOCATION: NUCLEAR.
CC	-I- SIMILARITY: BELONGS TO THE "DEAD" BOX FAMILY HELICASE. DEAH
CC	SUBFAMILY.
CC	-I- SIMILARITY: CONTAINS 2 DBRM (DOUBLE-STRANDED RNA-BINDING) DOMAINS.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC	or send an email to license@isb-sib.ch).
CC	-----
DR	EMBL; U91922; AAC05725.1; -
DR	EMBL; AF023530; AAC05301.1; -
DR	EMBL; U92080; AAB72087.1; -
DR	MGD; MGI:108177; Ddx9.
DR	InterPro: IPR001410; DEAD.
DR	InterPro: IPR002464; DEAH_ATP_helicase.
DR	InterPro: IPR001159; DS_RBD.
DR	InterPro: IPR001650; Helicase_C.
DR	Pfam: PF00035; dsrm; 2.
DR	Pfam: PF000271; helicase_C; 1.
DR	SMART; SM00487; DEXDC; 1.
DR	SMART; SM00358; DSRM; 2.
DR	SMART; SM00490; HELICC; 1.
DR	PROSITE: PS00690; DEAH_ATP_HELICASE; 1.
DR	PROSITE: PS50137; DS_RBD; 2.
KW	Helicase; RNA-binding; DNA-binding; Repeat; Nuclear protein;
KW	ATP-binding.
FT	DOMAIN 3 71 DRBM 1.
FT	DOMAIN 182 254 DRBM 2.
FT	NP_BIND 413 420 ATP (POTENTIAL).
FT	SITE 513 516 DEAH BOX.
FT	DOMAIN 588 597 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT	DOMAIN 1171 1380 ARG/GLY/SER/TYR-RICH.
FT	CONFLICT 46 46 R -> A (IN REF. 1; AAC05301).
FT	CONFLICT 674 674 N -> M (IN REF. 2).
FT	CONFLICT 748 748 T -> I (IN REF. 2).
FT	CONFLICT 831 831 I -> V (IN REF. 2).
SO	SEQUENCE 1380 AA; 149582 MW; 0F23755C291D2FE1 CRC64;

Query Match	21.2%; Score 180; DB 1; Length 1380;
Best Local Similarity	41.8%; Pred. NO. 9.9e-06;
Matches 59; Conservative	6; Mismatches 52; Indels 24; Gaps 9;
OY	22 GRLGSDLDTFGRVHGNLVAGIERACPRGYPLGTASIGGEVGARLCGRAGVGS-STYGY 80
DB	1216 GGFSGGGGGFGSGGGG-FGG---GGGGFSG--GGGGGGGGGGGGGGGGFGGNG 1268
OY	81 PSMGTPYGGGGYGGYGGYGGYGGDGRGASVGGYPCYGYG-----YPPSGYGGYGGSYGG 135
DB	1269 GGYGYGGGGYGG-GGGGGYGGGGSGYGG--GGYGGGGGYSISPNSIRGNVGGGGGGRGG 1325
OY	136 SYGSGSYTPN-----VRASAG 151
DB	1326 SGGG---YRNNGGDYRGSSG 1343

RESULT 11	
KICO_XENLA	
ID KICO_XENLA	STANDARD: PRT; 486 AA.
AC P02537;	
DT 21-JUL-1986 (Rel. 01, Created)	
DT 01-NOV-1988 (Rel. 09, Last sequence update)	
DT 20-AUG-2001 (Rel. 40, Last annotation update)	
DE KERATIN 3, TYPE I CYTOSKELETAL 51 KDA (51 KDA CYTOKERATIN).	
OS Xenopus laevis (African clawed frog).	
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;	
OC Xenopodinae; Xenopus.	

OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99005622; PubMed=2458965;
 RA Hoffmann W., Steiner S., Koenigsstorfer A.;
 RT "Amino acid sequence microheterogeneities of a type I cytokeratin of
 RT Mr 51,000 from *Xenopus laevis* epidermis.";
 RL FEBS Lett. 237:178-182(1988).
 RN [2]
 RP SEQUENCE OF 339-386 FROM N.A.
 RX MEDLINE=84261417; PubMed=6204859;
 RA Hoffmann W., Franz J.K.;
 RT "Amino acid sequence of the carboxy-terminal part of an acidic type I
 RT cytokeratin of molecular weight 51 000 from *Xenopus laevis* epidermis
 RL as predicted from the cDNA sequence.";
 RL EMBL J. 3:1301-1306(1984).
 CC -1- SUBUNIT: HETEROETRAMER OF TWO TYPE I AND TWO TYPE II KERATINS.
 CC -1- MISCELLANEOUS: THERE ARE TWO TYPES OF CYTOSKELETAL AND
 CC MICROFIBRILLAR KERATIN, I (ACIDIC) AND II (NEUTRAL TO BASIC)
 CC (40-55 AND 56-70 KILODALTONS, RESPECTIVELY).
 CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: Y00968; CAA68783.1; -;
 DR PIR: S01327; KRXL.
 DR InterPro: IPR001664; IF.
 DR InterPro: IPR002957; Keratin_I.
 DR Pfam: PF00038; filament; 1.
 DR PRINTS: PR01248; TYPE1KERATIN.
 DR PROSITE: PS00226; IF; 1.
 KW Intermediate filament; Coiled coil; Keratin.
 FT DOMAIN 1 125 HEAD.
 FT DOMAIN 126 438 ROD.
 FT DOMAIN 439 486 TAIL.
 FT DOMAIN 126 161 COIL_1A.
 FT DOMAIN 162 184 LINKER_1.
 FT DOMAIN 185 276 COIL_1B.
 FT DOMAIN 277 299 LINKER_12.
 FT DOMAIN 300 438 COIL_2.
 FT SITE 380 380 STUTTER.
 FT CONFLICT 352 352 R -> T (IN REF. 2).
 SQ SEQUENCE 486 AA; 51889 MW; 9B6D026C597109C5 CRC64;

 Query Match 21.0%; Score 179; DB 1; Length 486;
 Best Local Similarity 40.4%; Pred. No. 4.9e-06;
 Matches 44; Conservative 8; Mismatches 45; Indels 12; Gaps 4;

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93197168; PubMed=8451200;
 RA Good P.J., Rebert M.L., David I.B.;
 RT "Three new members of the RNP protein family in *Xenopus*.";
 RL Nucleic Acids Res. 21:999-1006(1993).
 CC -1- SUBCELLULAR LOCATION: NUCLEAR; COMPONENT OF RIBONUCLEOSOMES (BY
 CC SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE A/B GROUP OF hnRNP, WHICH ARE BASIC AND
 CC GLY-RICH PROTEINS.
 CC -1- SIMILARITY: CONTAINS 2 RNA RECOGNITION MOTIFS (RRM).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: L02957; AAA49950.1; -;
 DR HSSP: P09651; 10P1.
 DR InterPro: IPR000504; RRM.
 DR Pfam: PF00076; rrm; 2.
 DR SMART: SM00360; RRM; 2.
 DR PROSITE: PS0102; RRM; 2.
 DR PROSITE: PS00030; RRM_RNP_1; 2.
 KW Nuclear protein; RNA-binding; Ribonucleoprotein; Repeat.
 FT DOMAIN 27 110 RNA-BINDING (RRM) 1.
 FT DOMAIN 118 206 RNA-BINDING (RRM) 2.
 FT DOMAIN 207 385 GLY-RICH.
 SQ SEQUENCE 385 AA; 39438 MW; 96895F1B997C9E9D CRC64;

 Query Match 20.9%; Score 178; DB 1; Length 385;
 Best Local Similarity 32.5%; Pred. No. 4.8e-06;
 Matches 55; Conservative 12; Mismatches 61; Indels 46; Gaps 5;

RL J. Biol. Chem. 269:31457-31465(1994).
 CC -1- FUNCTION: MAY REGULATE MRNA TRANSLATION AND STABILITY. IT BINDS
 CC TO POLY(A) AND POLY(U) REGIONS OF RNA. THIS BINDING IS INHIBITED
 CC WHEN THE PROTEIN IS PHOSPHORYLATED.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC AND NUCLEAR. AT THE ONSET OF THE
 CC CRYPTOPHASIC PHASE, WHEN TRANSCRIPTION IS BLOCKED, IT ACCUMULATES
 CC IN THE CYTOPLASM.
 CC -1- PTM: EXTENSIVELY PHOSPHORYLATED ON TYROSINE RESIDUES.
 CC -1- SIMILARITY: BELONGS TO THE A/B GROUP OF HNRNP, WHICH ARE BASIC AND
 CC GLY-RICH PROTEINS.
 DR HSP, P09651: 10P1.
 DR InterPro: IPR00504; RRM.
 DR PROSITE: PS0102; RRM; PARTIAL.
 DR PROSITE: PS00030; RRM; RNP_1; PARTIAL.
 KM RNA-binding: Ribonucleoprotein; Nuclear protein; Phosphorylation;
 KM Methylation.
 FT NON_TER 1 1
 FT NON_CONS 15 16
 FT DOMAIN 32 >48
 FT NON_CONS 48 49
 FT NON_CONS 58 59
 FT NON_CONS 73 74
 FT NON_CONS 83 84
 FT NON_CONS 96 97
 FT DOMAIN 120 195
 FT MOD_RES 119 119
 FT MOD_RES 122 122
 FT MOD_RES 173 173
 FT MOD_RES 192 192
 FT NON_CONS 125 126
 FT NON_TER 195 195
 SO SEQUENCE 195 AA; 20604 MW; CA062DA8520B83DB CRC64;

Query Match 20.9%; Score 177.5; DB 1; Length 195;
 Best Local Similarity 43.4%; Pred. No. 3e-06;
 Matches 46; Conservative 7; Mismatches 30; Indels 23; Gaps 6;
 QY 42 IERAGPGYGLTASIGCEGARIGRAGVSSYGYGYGSMGYPYGYGYG----- 95
 DB 104 IKAATPDDAMMPMRNGRCGLGLG--AMVAPGSGFYG-----GGYGGGGYGDDA 154
 QY 96 YG--GYGYDQFCGSAVGYPGYGYPPSGYGGGSGYGSYG 139
 DB 155 YGAGGYDYGSGYGYGYS--GYEGY----GYNGGYGSGPANGG 194

RESULT 14
 ROAL_HUMAN STANDARD; PRT; 371 AA.
 AC P09651;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE HETEROGENEOUS NUCLEAR RIBONUCLEOPROTEIN A1 (HELIX-DESTABILIZING
 DE PROTEIN) (SINGLE-STRAND BINDING PROTEIN) (HNRNP CORE PROTEIN A1).
 GN HNRPA1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE OF 1-250 AND 303-371 FROM N.A.
 RC TISSUE=Liver;
 RA MEDLINE=89342435; PubMed=2760922;
 RA Blamont G., Buvoili M., Bassi M.T., Morandi C., Cobianchi F., Riva S.;
 RT "Isolation of an active gene encoding human hnrnp protein A1.
 RT Evidence for alternative splicing.";
 RL J. Mol. Biol. 207:491-503(1989).
 RN [2]
 RP SEQUENCE OF 1-250 AND 303-371 FROM N.A.
 RC TISSUE=Fibroblast;
 RX MEDLINE=88233978; PubMed=2836799;

RA Buvoili M., Blamont G., Ghetti A., Riva S., Bassi M.T., Morandi C.;
 RT "cDNA cloning of human hnrnp protein A1 reveals the existence of
 RT multiple mRNA isoforms.";
 RL Nucleic Acids Res. 16:3751-3770(1988).
 RN [3]
 RP SEQUENCE OF 1-250 AND 303-371 FROM N.A.
 RC TISSUE=Lung;
 RA Knudsen S.M., Leffers H.;
 RT Submitted (JUN-1994) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE OF 124-250 AND 303-371 FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=87053868; PubMed=3023065;
 RA Riva S., Morandi C., Tsoulfas P., Pandolfo M., Blamont G.,
 RA Merrill B., Williams K.R., Maltupa G., Beyreuther K., Werr H.,
 RA Heinrich B., Schaefer K.P.;
 RT "Mammalian single-stranded DNA binding protein UP I is derived from
 RT the hnrnp core protein A1.";
 RL EMBO J. 5:2267-2273(1986).
 RN [5]
 RP SEQUENCE OF 251-302 FROM N.A.
 RX MEDLINE=90214633; PubMed=1691095;
 RA Buvoili M., Cobianchi F., Bestagno M.G., Mangiarotti A., Bassi M.T.,
 RA Blamont G., Riva S.;
 RT "Alternative splicing in the human gene for the core protein A1
 RT generates another hnrnp protein.";
 RL EMBO J. 9:1229-1235(1990).
 RN [6]
 RP NUCLEAR LOCALIZATION DOMAIN.
 RX MEDLINE=95247808; PubMed=7730395;
 RA Sloml H., Dreyfuss G.;
 RT "A nuclear localization domain in the hnrnp A1 protein.";
 RL J. Cell Biol. 129:551-560(1995).
 RN [7]
 RP NUCLEAR LOCALIZATION DOMAIN, AND NUCLEAR EXPORT.
 RX MEDLINE=96067639; PubMed=8521471;
 RA Michael W.M., Choi M., Dreyfuss G.;
 RT "A nuclear export signal in hnrnp A1: a signal-mediated, temperature-
 RT dependent nuclear protein export pathway.";
 RL Cell 83:415-422(1995).
 RN [8]
 RP NUCLEAR LOCALIZATION DOMAIN.
 RX MEDLINE=95286702; PubMed=7769000;
 RA Weighardt F., Blamont G., Riva S.;
 RT "Nucleo-cytoplasmic distribution of human hnrnp proteins: a search
 RT for the targeting domains in hnrnp A1.";
 RL J. Cell Sci. 108:545-555(1995).
 RN [9]
 RP 3D-STRUCTURE MODELING OF 106-189.
 RX MEDLINE=91099515; PubMed=2176620;
 RA Ghetti A., Bolognesi M., Cobianchi F., Morandi C.;
 RT "Modeling by homology of RNA binding domain in A1 hnrnp protein.";
 RL FEBS Lett. 277:272-276(1990).
 RN [10]
 RP X-RAY CRYSTALLOGRAPHY (1.75 ANGSTROMS) OF 8-180.
 RX MEDLINE=97307256; PubMed=9164463;
 RA Shamoo Y., Krueger U., Rice L.M., Williams K.R., Steltz T.A.;
 RT "Crystal structure of the two RNA binding domains of human hnrnp A1
 RT at 1.75-A resolution.";
 RL Nat. Struct. Biol. 4:215-222(1997).
 RN [11]
 RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 6-181.
 RX MEDLINE=97277240; PubMed=9115444;
 RA Xu R.M., Jokhan L., Cheng X., Mayeda A., Krainer A.R.;
 RT "Crystal structure of human Upl, the domain of hnrnp A1 that contains
 RT two RNA-recognition motifs.";
 RL Structure 5:559-570(1997).
 CC -1- FUNCTION: INVOLVED IN THE PACKAGING OF PRE-MRNA INTO HNRNP
 CC PARTICLES, TRANSPORT OF POLY-A MRNA FROM THE NUCLEUS TO THE
 CC CYTOPLASM AND MAY MODULATE SPLICE SITE SELECTION.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR. SHUTTLES CONTINUOUSLY BETWEEN THE
 CC NUCLEUS AND THE CYTOPLASM ALONG WITH MRNA. COMPONENT OF
 CC RIBONUCLEOSOMES.

```

CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS, A1-A (SHOWN HERE) AND A1-B, ARE
CC PRODUCED BY ALTERNATIVE SPLICING. A1-A IS TWENTY TIMES MORE
CC ABUNDANT THAN A1-B.
CC -1- SIMILARITY: BELONGS TO THE A/B GROUP OF HNRNP, WHICH ARE BASIC AND
CC GLY-RICH PROTEINS.
CC -1- SIMILARITY: CONTAINS 2 RNA RECOGNITION MOTIFS (RRM).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; X12671; CA31191.1; -
DR EMBL; X06747; CA29922.1; ALT_SEQ.
DR EMBL; X04347; CA27874.1; -
DR EMBL; X79536; CA56072.1; -
DR PIR; S04617; S04617.
DR PIR; A24894; A24894.
DR PIR; S02061; S02061.
DR PDB; 1HA1; 15-MAY-97.
DR PDB; 1UP1; 17-SEP-97.
DR Aarhus/Ghent-2DPAGE; 207; NEPHGE.
DR Aarhus/Ghent-2DPAGE; 2114; NEPHGE.
DR Aarhus/Ghent-2DPAGE; 3612; NEPHGE.
DR MIM; 164017; -
DR InterPro: IPR000504; RRM.
DR Pfam; PF00076; rrm; 2.
DR SMART; SM00360; RRM; 2.
DR PROSITE; PS50102; RRM; 2.
DR PROSITE; PS00030; RRM_RNP_1; 2.
DR Nuclear protein; RNA-binding; Repeat; Ribonucleoprotein;
DR Methylation; Alternative splicing; 3D-structure.
FT INT_MET 0
FT DOMAIN 3 93 GLOBULAR A DOMAIN.
FT DOMAIN 94 184 GLOBULAR B DOMAIN.
FT DOMAIN 13 96 RNA-BINDING (RRM) 1.
FT DOMAIN 104 183 RNA-BINDING (RRM) 2.
FT DOMAIN 217 239 RNA-BINDING RGG-BOX.
FT DOMAIN 194 371 GLY-RICH.
FT DOMAIN 319 356 NUCLEAR TARGETING SEQUENCE (M9).
FT MOD_RES 193 193 METHYLATION (BY SIMILARITY).
FT VAAPPLIC 251 302 MISSING (IN ISOFORM A1-A).
FT MUTAGEN 325 325 G->A: NO NUCLEAR IMPORT NOR EXPORT.
FT MUTAGEN 326 326 P->A: NO NUCLEAR IMPORT NOR EXPORT.
FT MUTAGEN 333 334 GG->LL: NORMAL NUCLEAR IMPORT AND EXPORT.
FT CONFLICT 139 139 R->P (IN REF. 4).
SQ SEQUENCE 371 AA; 38715 MW; B3EFA5A1DB7C26 CRC64;

```

Query Match 20.8%; Score 177; DB 1; Length 371;
 Best Local Similarity 33.1%; Pred. No. 5.4e-06;
 Matches 47; Conservative 14; Mismatches 53; Indels 28; Gaps 4;

```

QY 14 ALTNAAARGLSDLPFGRAVHGNLYAGIERAPRGYPGLTASTIGCEVGARLGAGAGV 73
DB 186 ASASSSQGRSGS--GNFGGGRGGGFGGNDNFR-----GNFNSGGRGG 229
QY 74 SSYGIGYPSKGPYGGYGGYGGY-----GYGTYDQGRGSAVGYPIYGY-Y 121
DB 230 SRGGGGYSGSDYNGFENDGCGPGYSGSGYSGGCGYGNQSGYGGSGSYDSY 289
QY 122 PSYGCGGCGSYGSGSYGY 143
DB 290 NNGGGRGFGGSGSNFGGGSY 311

```

RESULT 15
 DDX9_BOVIN
 AC Q28141; STANDARD; PRT; 1287 AA.

```

DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE ATP-DEPENDENT RNA HELICASE A (NUCLEAR DNA HELICASE II) (NDH II)
DE (DEAD-BOX PROTEIN 9).
CN DDX9 OR NDH2.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Thymus;
RX MEDLINE=9533357; PubMed=7608213;
RA Zhang S., Maacke H., Grosse F.;
RT "Molecular cloning of the gene encoding nuclear DNA helicase II. A
RT bovine homologue of human RNA helicase A and Drosophila Mle
RT protein."
RL J. Biol. Chem. 270:16422-16427(1995).
CC -1- FUNCTION: UNWINDS DOUBLE-STRANDED DNA AND RNA IN A 3' TO 5'
CC DIRECTION. GENERATES MULTIPLE SECONDARY STRUCTURES THAT INFLUENCE
CC RNA-BINDING PROTEINS. MAY PLAY A ROLE IN X-LINKED GENE EXPRESSION.
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- SIMILARITY: BELONGS TO THE "DEAD" BOX FAMILY HELICASE. DEAD
CC SUBFAMILY.
CC -1- SIMILARITY: CONTAINS 2 DRBM (DOUBLE-STRANDED RNA-BINDING) DOMAINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; X82829; CA58036.1; -
DR InterPro: IPR001410; DEAD.
DR InterPro: IPR002464; DEAD_ATP_helcse.
DR InterPro: IPR001159; DS_RBD.
DR InterPro: IPR001650; Helicase_C.
DR Pfam; PF00035; dsrm; 2.
DR Pfam; PF00271; helicase_C; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00358; DSRM; 2.
DR SMART; SM00490; HELIC; 1.
DR PROSITE; PS00690; DEAD_ATP_HELICASE; 1.
DR PROSITE; PS50137; DS_RBD; 2.
DR Helicase; RNA-binding; DNA-binding; Repeat; Nuclear protein;
KW ATP-binding.
KW DOMAIN 3 71 DRBM 1.
FT DOMAIN 177 249 DRBM 2.
FT NP_BIND 408 415 ATP (BY SIMILARITY).
FT SITE 508 511 DEAD BOX.
FT DOMAIN 583 592 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT DOMAIN 1167 1287 ARG/GLY/TYR-RICH.
SQ SEQUENCE 1287 AA; 141943 MW; DC908095AB683ED4 CRC64;

```

Query Match 20.7%; Score 176.5; DB 1; Length 1287;
 Best Local Similarity 37.9%; Pred. No. 1.6e-05;
 Matches 53; Conservative 1; Mismatches 45; Indels 41; Gaps 5;

```

QY 37 NLVYGIERAPRGYPGLTASTIGCEVGARLGAGAGVSSY-GYGPYSGYGGYGG 95
DB 1144 NLMITGTRVYGGPPPKMARIDNDSGYTRRG-----SSYSGGYGLAGYTGCGGGGG 1197
QY 96 YGSGYGYDQG-----FGSAY-----GGYPYGYYPYSGYGG 128
DB 1198 YGGRGYSGGSGGSGNFRSGYVGGGGGCGFPGGLSRGGRGSGDYGESGCG 1257
QY 129 YGGSYG-----GSYGGSY 141

```

'Thu Jan 17 08:02:54 2002

us-09-554-547-16.rsp

Page 11

Db 1258 YRSGGF0RGGGRGCGY 1277

Search completed: January 16, 2002, 14:58:40
Job time: 192 sec

THIS PAGE BLANK (USPTO)

1	207.5	24.4	423	10	Q9FM47	Q9FM47 arabidopsis
2	206.5	24.3	159	5	Q9VE11	Q9VE11 drosophila
3	198	23.3	464	10	Q9FMK8	Q9FMK8 cryza sativa
4	195.5	23.0	245	5	Q9VZ59	Q9VZ59 drosophila
5	193	22.7	159	11	Q08884	Q08884 mus musculus
6	192	22.6	458	5	Q9N3U1	Q9N3U1 caenorhabdl
7	191.5	22.5	146	5	Q9VFT9	Q9VFT9 drosophila
8	191	22.4	141	5	Q9VNN3	Q9VNN3 drosophila
9	190.5	22.4	259	5	Q02049	Q02049 caenorhabdl
10	187	22.0	263	5	Q40056	Q40056 candida alb
11	185.5	21.8	250	5	Q9VX64	Q9VX64 drosophila
12	184.5	21.7	440	5	Q9V6M1	Q9V6M1 drosophila
13	183	21.5	173	5	Q9VX65	Q9VX65 drosophila
14	182.5	21.4	142	5	Q9U517	Q9U517 manduca sex
15	181.5	21.3	305	11	Q9CX86	Q9CX86 mus musculus
16	180	21.2	358	10	Q9FHL2	Q9FHL2 arabidopsis
17	180	21.2	360	5	Q16965	Q16965 araneus dia
18	179.5	21.1	261	10	Q9ZFM2	Q9ZFM2 cucumis sat
19	179	21.0	268	10	Q9ZFM2	Q9ZFM2 arabidopsis

20	178.5	21.0	289	10	09G09	09G09	arabidopsis
21	178.5	21.0	303	5	017201	017201	bombyx mori
22	177	20.8	691	5	09B103	09B103	dolomedes t
23	176	20.7	196	5	02A346	02A346	penicillium
24	175.5	20.6	197	10	09APF3	09APF3	penicillium
25	175.5	20.6	309	10	09FMR1	09FMR1	arabidopsis
26	175	20.6	154	10	09SL23	09SL23	arabidopsis
27	175	20.6	181	5	09VNM4	09VNM4	droserophila
28	174.5	20.5	208	10	09ZRV2	09ZRV2	cicer arile
29	174.5	20.5	988	5	017434	017434	nephila cla
30	174	20.4	347	10	001916	001916	phytophtho
31	174	20.4	576	5	09VX72	09VX72	droserophila
32	173	20.3	193	11	070592	070592	rattus norv
33	172.5	20.3	255	10	09SH12	09SH12	arabidopsis
34	172	20.2	317	5	09W020	09W020	droserophila
35	172	20.2	404	10	09LKA4	09LKA4	arabidopsis
36	171.5	20.2	167	10	024184	024184	oryza sativ
37	171.5	20.2	165	10	045158	045158	triticum ae
38	171.5	20.2	287	5	017200	017200	bombyx mori
39	171	20.1	126	5	044355	044355	aegipecten
40	170	20.0	136	5	044354	044354	aegipecten
41	169.5	19.9	133	5	09UAV7	09UAV7	caenorhabd1
42	169.5	19.9	173	10	040057	040052	hordium vul
43	169.5	19.9	2639	5	076786	076786	atheraea p
44	169	19.9	127	11	070370	070370	mus muscul
45	169	19.9	209	5	09VX73	09VX73	droserophila

ALIGNMENTS

RESULT	1	
09Fma47		
ID	PRELIMINARY;	PRT; 423 AA.
AC	09Fma47	
DT	01-MAR-2001 (TREMBLrel. 16, Created)	
DT	01-MAR-2001 (TREMBLrel. 16, Last sequence update)	
DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)	
DE	SIMILARITY TO RNA BINDING PROTEIN.	
OS	Arabidopsis thaliana (Mouse-ear cress).	
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae	
OC	eucosids. II; Brassicales; Brassicaceae; Arabidopsids.	
OX	NCBI_TaxID=3702;	
RA	11	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=COLUMBIA.	
RX	MEDLINE=98290546; PubMed=9628582;	
RA	Sato S., Kaneko T., Kotani H., Nakamura Y., Asamizu E., Miyajima N.,	
RA	Tabata S.;	
RT	"Structural analysis of Arabidopsis thaliana chromosome 5. IV.	
RT	Sequence features of the regions of 1,456,315 bp covered by nineteen	
RL	physically assigned P1 and TAC clones."	
RL	DNA Res. 5:41-54 (1998).	
DR	EMBL; AB009052; BAE08520.1; .	
DR	InterPro; IPR002952; Eggshell1.	
DR	InterPro; IPR00504; RRM.	
DR	Pfam; PF00076; rrm. 2.	
DR	PRINTS; PR0128; EGGSHLL.	
DR	SMART; SM00360; RRM. 2.	
DR	PROSITE; PS50102; RRM. 2.	
DR	PROSITE; PS00030; RRM_RNP_1; 1.	
Q0	SEQUENCE 423 AA; 42408 MW; DCFB04E24CF10002 CRC64;	

Query Match	24.48;	Score 207.5;	DB 10;	Length 423;
Best Local Similarity	40.58;	Pred. No. 6.5e-10;		
Matches 60; Conservative	4;	Mismatches 51;	Indels 33;	Gaps 7;

OY 23 RLGSDDLTPGRVHGNLTLAGIEAGPRGIPGLTISIGGEFVARLGGKAGVGYSSSYGYTPS 82
| | : | | | | | | | |
Db 220 RFGSDRSNFEGCGYGDIYGGCHGGGYYGGRPYKSGGYGGGRSGCYGYGGEFFGYYGGG 279

QY 83 WGYPGYGGY-----GGYGGY--GGYDGFSGSNAGYGGPG-VYGYVY----- 121
 Db 280 YG-----GGVGPRIPEPALCYSGRYGGGGGYNMGSGGGGGYGGPGDMYGGSYGEPBG 336
 QY 122 ----PSG-YGGGYG----GSYGGSYGGS 140
 Db 337 GYGPGSGYGGYGGSGIGYGGGMGA 364

RESULT 2

09VE11 ID 09VE11 PRELIMINARY: PRT: 159 AA.
 AC 09VE11;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE E0G91 PROTEIN.
 CN E0G91 OR CG7539.
 OS Drosophila melanogaster (fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 CX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Branton R.C., Rogers V.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abell J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Beron P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Broksstein P., Brotlier P.,
 RA Butris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherly J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Itoigawa C.,
 RA Jallat M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Maltel B., McIntosh T.C., Moleod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Mount S.M., Nelson K.A., Nixon K., Nusskern D.R., Paclib J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Smith H.,
 RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Sheth T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svaydas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zhang X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003718; AARS5440.1;
 DR FLYBase: FBgn004354; Edg91.
 DR InterPro: IPR002952; EggsHELL.
 DR PRINTS: PRO1228; EGGSHELL.
 SO SEQUENCE 159 AA; 15261 MW; 266216A2A39AA06F CRC64;

Best Local Similarity 42.0%; Pred. No. 2,6e-10;
 Matches 58; Conservative 7; Mismatches 34; Indels 39; Gaps 10;
 QY 38 LYAGIERAGPRGCGYRGLTAS-IGGEVARGALRGAGVSS-----YGYGP---SMGYPY 87
 Db 14 LIAGGGAAPYKTEESTLGLGGGFGSGVLSAGIGVGGGLSGGCGGYRGTASGTP- 72
 QY 88 GGY-GGYGGYGGYGGYDGFSGAY---GGYPGY-----YGYVPSG-----YGGYGGGS 132
 Db 73 GYGGGYSGYNGYGG--SGFGGYRPGGYSGFGRPHNHGGYRPGGSGYNNHGGSGYCGH 130
 QY 133 Y-----GSGYGG 139
 Db 131 YSOSQYSNGYGGGYYGG 148

RESULT 3

09FWK8 ID 09FWK8 PRELIMINARY: PRT: 464 AA.
 AC 09FWK8;
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE PUTATIVE RNA BINDING PROTEIN.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Eriobotryaceae; Oryzae; Oryza.
 CX NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. NIPPONBARE;
 RX Buiji C.R., Yuan Q., Moffat K.S., Hill J.N., Jenkins C.N., Burr P.C.,
 RA Hsiao J., Zisman V., Pai G., Bowman C.L., Fujii C.Y., VanKen S.E.,
 RA Bowman C.L., Craven B., Utterback T.R., Khalak H., Feldolym T.V.,
 RA Quackenbush J., White O., Salzberg S.L., Fraser C.M.;
 RT "Oryza sativa chromosome 10 BAC OSJNBa007916 genomic sequence.";
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC026815; ACG1903.1;
 DR InterPro: IPR000504; RRM.
 DR Pfam: PF00076; rrm; 2.
 DR SMART: SM00360; RRM; 2.
 DR PROSITE: PS50102; RRM; 2.
 DR PROSITE: PS00030; RRM_RNP_1; UNKNOWN_1.
 SO SEQUENCE 464 AA; 45621 MW; CFC03DD3045E4B92 CRC64;

Query Match 23.3%; Score 198; DB 10; Length 464;
 Best Local Similarity 30.8%; Pred. No. 4,6e-09;
 Matches 60; Conservative 11; Mismatches 48; Indels 76; Gaps 9;

QY 21 AGRGSDUDTFRHGNLYAGIERAGPRGPGTLTASIGEGV----- 62
 Db 263 AGGYGAGVGYGYGANGTAYGNTGAGGFGVPGTGYGNPNAGPSGYQGCPPCANR 322
 QY 63 ARLGGR--GYYVSSYG--YGYPSM-----GYPYGGYG-- 92
 Db 323 GPMVGQAAPSGYCTGSGYGNAGYAAAMNNSAGCNAPTSGAAGAGTGYGSGYGGYGGDA 382
 QY 93 ----YGGYGGYGGYDQFGS-AYGGYPGYGY-----YPSGY-----GGYGGG-- 132
 Db 383 SYGNHGGYGGYGGGDDGAGNPAAGGSGYAGYSGNGSGYPPNAMPDSGGGFGASVN 442
 QY 133 -----YGGSYGG 139
 Db 443 GVSQGQSNVSGSYGG 457

Query Match

24.3%; Score 206.5; DB 5; Length 159;

RESULT 4
 09VZ59 ID 09VZ59 PRELIMINARY: PRT: 245 AA.
 AC 09VZ59;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)

DT 01-MAV-2000 (Tremblay et al. 2000, last sequence update)

DT 01-MAV-2000 (Tremblay et al. 2000, last sequence update)

DE CC2157 PROTEIN.

GN Drosophila melanogaster (fruit fly).

OS Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Pnychoidea; Drosophilidae; Drosophila.

OX NCBI_TaxID=7227;

XX [1]

RP SEQUENCE FROM N.A.

RP STRAIN=BERKELEY.

RX MEDLINE=20196006; PubMed=10731132;

RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Amanalides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,

RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RA Sulten G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,

RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,

RA Non K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,

RA Abtill J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baltow D.,

RA Bailew R.M., Bass A., Bayendale J., Bayraktaroglu L., Beasley E.M.,

RA Beeson K.Y., Benos P.V., Brentan B.P., Bhargava D., Bolshakov S.,

RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,

RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,

RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,

RA de Paulos B., Delcher A., Deng Z., Duan A.D., Dew I., Dietz S.M.,

RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrieres S., Fleischmann W.,

RA Folsler C., Gabrielian A.E., Gang N.S., Gelbart W.M., Glasser K.,

RA Glodde A., Gong F., Gorell J.H., Gu Z., Guan P., Harris M.,

RA Harits N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,

RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,

RA Jallat M., Kalish F., Karpen G.H., Ke Z., Kennison J.A., Kethum K.A.,

RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

RA Laslo P., Lai Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,

RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,

RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,

RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,

RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Paolel J.M.,

RA Palczak M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,

RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,

RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,

RA Spter E., Spradling A.C., Stapleton M., Strong R., Sun E.,

RA Svatek R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,

RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weisenbach J.,

RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.-A.,

RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,

RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,

RA Gdbbs R.A., Myers E.W., Rubin G.M., Venter J.C.;

RT "The genome sequence of Drosophila melanogaster.";

RL Science 287:2185-2195(2000).

DR EMBL: AEO03484; AAF47968.1; -

DR FlyBase; FBgn0030244; CG2157.

QC SEQUENCE 245 AA: 32982 MW: 8386620640035C CRC64:

Query Match	23.0%	Score 195.5	DB 5	Length 245
Best Local Similarity	31.0%	Pred. No. 3.6e-09		
Matches	61	Conservative	14	Mismatches 67; Indels 55; Gaps
QY	1	MKAFFVLSLSTAAITLNARAGRIGSDLDI	-----	30
Db	1	MKCFPI--LIATSLVLSVAQDAADSDLEQVASASAPDAVAVEDQVSAAVSAPLQDKQ		58
QY	31	-----FGRVHCILYAGIERAERPRGTRPGLTASIGGEVAKRLGGRRAGVYSSTGYGT		80
Db	59	EKRRYVGSGYGGAGCYGAGYGGAGYAGGAGGAGSGLGAGYGGAGVSGYGVDSVYNDY		118
QY	81	PSMGYPYGGYGC--YGGYGGYGGDGGFGSAYGYPGYYYYPSSGYYGGYGGSGYSG		138
Db	119	SS--RALGGLDGAANVAGTCGTSLGGYSALGGYSALC-----GYGSLG--YGGSTN		170
QY	139	G---SYTYPPNVRASAGA		152

Db	171	SNPYALSTYNSRLGACA	187
RESULT	5		
ID	008884	PRELIMINARY;	PRT; 159 AA.
AC	008884		
DT	01-JUL-1997	(TREMBLrel. 04, Created)	
DT	01-JUL-1997	(TREMBLrel. 04, Last sequence update)	
DT	01-OCT-2000	(TREMBLrel. 15, Last annotation update)	
DE	HIGH-GLYCINE TYROSINE KERATIN TYPE II.4.		
GN	KRTAP6-2.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Cranita; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=SKIN;		
RA	Aoki N.;		
RT	"High-glycine / tyrosine keratin.";		
RL	Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.		
DR	EMBL: D89902; BAA20281.1; "		
DR	MED: MGI:1330280; Krtap6-2.		
DR	Keratin.		
SO	SEQUENCE 159 AA; 15262 MW; 2925AEB08A777BB9 CRC64;		
Query Match	22.7%;	Score 193;	DB 11; Length 159;
Best Local Similarity	46.2%;	Pred. No. 3.5e-09;	
Matches	55; Conservative	5; Mismatches	45; Indels 14; Gaps 9;
Qy	31	FGRVHGNLYAGIERAPPGYPGLTASIGCEVGARLGRAGVGV-SSYGYGYP-WGYPYG 88	
Db	16	YGSGYGSFY-----GCGSSGYGCGGYSGYGCYGGSGSGYGCYGGSGYGCYGC 69	
Qy	89	GYG-GYG-GYG-GYG-GYG-GYDQFGSAYG-GTPGYIGYTPSGYGGYGGSGSGSGSY 141	
Db	70	SGYGGYSGYSGYGCYGGSGYGCYGGYGCYGGYGCYGGYGCYGCYGCYGCYGCYGCY 128	
RESULT	6		
ID	Q9N301	PRELIMINARY;	PRT; 458 AA.
AC	Q9N301		
DT	01-OCT-2000	(TREMBLrel. 15, Created)	
DT	01-OCT-2000	(TREMBLrel. 15, Last sequence update)	
DT	01-JUN-2001	(TREMBLrel. 17, Last annotation update)	
DE	HYPOTHEITICAL PROTEIN Y47D7A.B.		
GN	Y47D7A.B.		
OS	Caenorhabditis elegans.		
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdilita; Rhabditidae;		
OC	Rhabdilitae; Peloderinae; Caenorhabditis.		
OX	NCBI_TaxID=6239;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=BRISTOL N2;		
RC	MEDLINE=99069613; PubMed=9851916;		
RA	None;		
RT	"Genome sequence of the nematode C. elegans: a platform for		
RT	investigating biology. The C. elegans Sequencing Consortium.";		
RL	Science 282:2012-2018(1998).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=BRISTOL N2;		
RA	Waterston R.;		
RL	Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.		
DR	EMBL: AC024790; AAF60630.1; "		
DR	InterPro: IPR002952; Eggshell.		
DR	InterPro: IPR000976; Wilms-tumour.		
DR	PRINTS: PR01228; EGGSHELL.		
DR	PRINTS: PR00049; WILMSTUMOUR.		

SO SEQUENCE 458 AA; 40183 MW; 3CF054AE11209FE6 CRC64;

Query Match 22.6%; Score 192; DB 5; Length 458;

Best Local Similarity 37.2%; Pred. No. 1.4e-08; Matches 55; Conservative 10; Mismatches 59; Indels 24; Gaps 6;

QY 22 GRLGSDLDTPGVRHGNL--YAGIERAPRGYPGLTASIGGEVGARLGRACGVSSYSG 79
 Db 227 GYAGAGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 286
 QY 80 YPSMGYPYGGYGGYGGY-----GGYGGYDQ---GFGSAYGGYPGYGYGYSSGYCG 128
 Db 287 ---GYACAGCGGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 337
 QY 129 YGSGYSG--YGGSYTPNVRASAGAAA 154
 Db 338 FGGSGGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 365

RESULT 7

Q9VFT9 PRELIMINARY; PRT; 146 AA.
 AC Q9VFT9; 01-MAY-2000 (TREMblrel. 13, Created)
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
 DT 01-MAY-2000 (TREMblrel. 13, Last annotation update)
 DE CG9269 PROTEIN.
 GN CG9269.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NCBI_Taxid=7227;
 [1]
 RN NCBI_Taxid=7227;
 RP SEQUENCE FROM N.A.
 RC STRAIN-BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
 Wan K.H., Doyle C., Baxter E.G., Helt J., Nelson C.R., Miklos G.L.G.,
 Abail J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,
 Butts K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 Durbin K.J., Evangelista C.C., Ferraz C., Ferriere S., Fleischmann W.,
 Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 Jaitai M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacלב J.M.,
 Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissenbach J.,
 Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,

RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 "The genome sequence of Drosophila melanogaster";
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003701; AAF54960.1;
 DR FlyBase: FBgn0038161; CG9269.
 SO SEQUENCE 146 AA; 14788 MW; 35DD14C33544BFA3 CRC64;

Query Match 22.5%; Score 191.5; DB 5; Length 146;
 Best Local Similarity 50.0%; Pred. No. 4.3e-09; Matches 46; Conservative 3; Mismatches 26; Indels 17; Gaps 7;

QY 65 LGRAGVSSYGYGYPGYPGYPGYPGYPGYPGYPGYPGYPGYPGYPGYPGYPG 124
 Db 46 LGGYGG-----YGSYGGFGGYPGYPGYPGYPGYPGYPGYPGYPGYPGYPGYPG 98
 QY 125 --YGG---GYGSGYGG---SYGSGYT--YPNV 146
 Db 99 NLGGSGGLAGGYGYPGYPGYPGYPGYPGYPGYPGYPGYPGYPGYPGYPGYPG 130

RESULT 8

Q9VNN3 PRELIMINARY; PRT; 141 AA.
 AC Q9VNN3; 01-MAY-2000 (TREMblrel. 13, Created)
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
 DE CG1597 PROTEIN.
 GN CG1597.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NCBI_Taxid=7227;
 [1]
 RN NCBI_Taxid=7227;
 RP SEQUENCE FROM N.A.
 RC STRAIN-BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
 Wan K.H., Doyle C., Baxter E.G., Helt J., Nelson C.R., Miklos G.L.G.,
 Abail J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,
 Butts K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 Durbin K.J., Evangelista C.C., Ferraz C., Ferriere S., Fleischmann W.,
 Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 Jaitai M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacלב J.M.,
 Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissenbach J.,
 Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,

01-MAY-2000 (TReMBLrel. 13. Created)
01-MAY-2000 (TReMBLrel. 13. Last sequence update)
01-JUN-2001 (TReMBLrel. 17. Last annotation update)
CG10597 PROTEIN.
CG10597.
Drosophila melanogaster (Fruit fly).
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
NCBI TaxID=7227;
[1]
SEQUENCE FROM N.A.
RP STRAIN-BERKELEY.
RC MEDLINE-20196006; PubMed-10731132:
RX MEDLINE-20196006; PubMed-10731132:
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
Ramey A., Smiley A., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Mortman J.R., Yandell M.D., Zheng Q., Chen L.X.,
RA Branton R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abail J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,
RA Bokorova D., Bolcham M.R., Bouck J., Brooksstein P., Brothier P.,
RA Butts K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbini K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek C., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Houlton D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jaitani M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Krell C., Kravitz S., Kulp D., Lai Z.,
RA Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Meltel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Mekulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacle J.M.,
RA Palazzolo M., Plutman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodgett T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster";
RL Science 287:2185-2195(2000).
DR EMBL: AEO03504; AAF48714.1; -
DR FlyBase: FBgn0030832; CG10597.
DR InterPro: IPR002952; Eggshell.
DR PRINTS: P001228; EGGSHLL.
SQ SEQUENCE 250 AA: 23045 MW: 9A7DE402E0DF038E CRC64;

Db	177	AGGAGC--GGAGACRGAGAGCGAGAGAGAGAGMAGMAGNAEPDITYNNHASGAGAGATG	233
Oy	136	SYGSSYTPNVNRSASG	151
	:		
Db	236	AAGG-----GSRGAG	246
RESULT	12		
Q9V6M1	PRELIMINARY;	PRT;	440 AA.
AC	Q9V6M1;		
DT	01-MAY-2000 (TREMBLrel. 13, Created)		
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)		
DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)		
DE	CG4663 PROTEIN.		
GN	CG4663.		
OS	Drosophila melanogaster (Fruit fly).		
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;		
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;		
OC	Ephydroidea; Drosophilidae; Drosophila.		
OX	NCBI_TaxID=7227;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=BERKELEY.		
RX	MEDLINE=20196006; PubMed=10731132;		
RA	Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,		
RA	Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,		
RA	Sutton R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,		
RA	Geiton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,		
RA	Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,		
RA	Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,		
RA	Abtill J.F., Agbayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,		
RA	Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,		
RA	Besson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,		
RA	Botkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,		
RA	Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,		
RA	Cherry J.M., Cawley S., Dahlke C., Davensport L.B., Davies P.,		
RA	de Pablos B., Delcher A., Deng Z., Dey A.D., Dew I., Dietz S.M.,		
RA	Dooson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,		
RA	Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,		
RA	Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,		
RA	Glodet A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,		
RA	Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,		
RA	Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegam C.,		
RA	Jatall M., Kalush E., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,		
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,		
RA	Liao K.P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,		
RA	Li X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,		
RA	MeKulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,		
RA	Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,		
RA	Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacheb J.M.,		
RA	Palazzolo M., Pittman G.S., Pan S., Pollard J., Puti V., Reese M.G.,		
RA	Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,		
RA	Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,		
RA	Spier E., Spreading A.C., Stapleton M., Strong R., Sun E.,		
RA	Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,		
RA	Wang Z.-T., Wassarman D.A., Weislock G.M., Weissbach J.,		
RA	Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,		
RA	Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,		
RA	Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,		
RA	Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;		
RT	"The genome sequence of Drosophila melanogaster.";		
RL	Science 287;2185-2195(2000).		
DR	EMBL; AE003819; AAF58402.1; -		
DR	FLYBase; FBgn0033812; CG4663.		
DR	InterPro; IPR001452; SH3.		
DR	Pfam; PFO0018; SH3; 1.		
DR	SMART; SM00326; SH3; 1.		
DR	PROSITE; PSS0002; SH3; 1.		
SO	SEQUENCE	440 AA; 46660 MW; 3ABEC719DDE9G50D CRC64;	

Query Match 21.78; Score 184.5; DB 5; Length 440;

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☒ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☐ FADED TEXT OR DRAWING
- ☒ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☒ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.

THIS PAGE BLANK (USPTO)